Thesis

Non-verbal expression perception and mental state attribution by third parties

HYNIEWSKA, Sylwia

Abstract

Understanding the internal states of others is essential in social exchanges. The aim of the presented thesis is to provide a deeper understanding of the impact nonverbal cues have on the perception of internal states, namely of emotions and associated cognitive appraisals. First, we explored naturalistic behaviour from a hidden camera, described with technical coding systems, FACS for perceived facial muscle movements and a coding scheme defined by ourselves for the hand, arm and torso movements. Participants were asked to judge observed persons' internal states. These descriptions of behaviours and perceptive judgments allow us to make a link between internal state attributions and concrete physical expressions. Second, a novel method was used for expression exploration by transposing naturalistic behaviours to a virtual agent, Greta, which enables a fine tuning of expressions. In order to improve the synchronisation of behaviours the Multimodal Sequential Expression model was created for the Greta agent. Complex expressions were manipulated, one cue at a time, and expression were judged by participants, who were asked to attribute internal states to the agent. Results support the componential approaches to expression, in [...]

Reference

HYNIEWSKA, Sylwia. *Non-verbal expression perception and mental state attribution by third parties*. Thèse de doctorat : Univ. Genève, 2013, no. FPSE 525

Sous la direction de Catherine Pelachaud et Susanne Kaiser

Non-verbal expression perception and mental state attribution by third parties

THESE

Présentée à Télécom ParisTech
pour obtenir le grade de Docteur en Informatique
N° 2013ENST0012

et à la Faculté de psychologie et des sciences de l'éducation de l'Université de Genève
pour obtenir le grade de Docteur en Psychologie
N° 525

par
Sylwia Julia HYNIEWSKA

de
Varsovie, Pologne

PhD thesis defended on 14th of March 2013

To my father who, since my toddler years, has always found time to sit and think with me. Dad, thank you for stories of robots, political conspiracy theories, science discoveries and for taking all my questioning seriously.

Table of content

Abstract

Understanding the internal states of others is essential in social exchanges. The aim of the presented thesis is to provide a deeper understanding of the impact nonverbal cues have on the perception of internal states, namely of emotions and associated cognitive appraisals. First, we explored naturalistic behaviour from a hidden camera, described with technical coding systems, FACS for perceived facial muscle movements and a coding scheme defined by ourselves for the hand, arm and torso movements. Participants were asked to judge observed persons' internal states. These descriptions of behaviours and perceptive judgments allow us to make a link between internal state attributions and concrete physical expressions. Second, a novel method was used for expression exploration by transposing naturalistic behaviours to a virtual agent, Greta, which enables a fine tuning of expressions. In order to improve the synchronisation of behaviours the Multimodal Sequential Expression model was created for the Greta agent. Complex expressions were manipulated, one cue at a time, and expression were judged by participants, who were asked to attribute internal states to the agent. Results support the componential approaches to expression, in which particular cues are considered meaningful.

Preface

Our whole cubic capacity is sensibly alive; and each morsel of it contributes its pulsations of feeling, dim or sharp, pleasant, painful, or dubious, to that sense of personality that every one of us unfailingly carries with him.

James, 1884

Foreword

Processes underlying emotion recognition have been studied extensively over the past decades, yet our understanding remains incomplete – however essential the actual task of decoding the mental state of the individuals who surround us may be . It is important for personal and functional reasons, such as detection of threat or acknowledgment of a loss, but also for the well being of a group of individuals. Interpretation of others' affects and motives is also important in short term social exchanges (e.g. to make sense of ambiguous social situations; Manstead & Fischer, 2001), as well as in the long term ones, starting from building healthy child-caregiver attachment (e.g. Bowlby, 1969) to adult social relations (Feldman Barrett & Salovey, 2002; Salovey & Mayer, 1990; Vrtička & Vuilleumier, 2012). Altered emotion perception is associated with a variety of neuropsychological conditions and other high-risk populations (Davidson, Putnam & Larson, 2000; Kohler, Turner, Gur & Gur, 2004; Demenescu, Kortekaas, den Boer & Aleman, 2010; Brosnan, Ashwin & Gamble, 2012) and investigating recognition of emotional facial expression deficits has potential implications for diagnosis of different pathologies (Bediou et al., 2012).

Despite the great use that a clearer comprehension of emotions would have in the applications mentioned above, and in many others, the current understanding of emotions is still far from enabling advanced affective analysis. Scholars have not agreed on a standard definition of emotion, and there is a long history of debates surrounding this term. Few researchers involved in emotion studies would not agree with Fehr and Russell's conclusion (1984) that "*Everyone knows what an emotion is, until asked to give a definition. Then it seems that no one knows*".

Although feeling often plays an important role in the various definitions of the term "emotion", particularly for laypeople, this is probably not the only phenomenon that leads to the formation of the concept of emotion (Frijda, 2000). Indeed, emotions can be seen as exclusively purely bodily feelings, action tendencies or cognitive evaluations, or, in a majority of cases, as a product of many such components.

All of these components build up to a global understanding of emotion, and the way individuals perceive, experience and express them. There is still an open discussion about when and how emotions arise as a result of these different components, e.g. the implication of cognitive processes triggered by everyday life experience, such as confronting some new events or reacting to others' behaviour.

In the last decades, psychologists have been interested in investigating the cognitive process in the elicitation of emotion. One of the main difficulties arising in this exploration is the lack of direct and objective measures of emotions and their components. It is indeed hard to access emotions by any other means than verbal self-reports, which are typically used in studies (see Frijda, Kuipers & ter Schure, 1989; Fredrickson, 2000; Siemer, Maus & Gross, 2007). Furthermore, individuals often find it difficult to observe, assess and describe verbally their own or others' cognitive components, such as appraisals of happening events. Nevertheless, some theories (e.g. Scherer, 2000; Scherer & Ellgring, 2007) claim a direct link between precise patterns of appraisals and emotions and nonverbal cues, such as facial expressions. Identifying which nonverbal cues are associated with emotions and appraisals – whether at the expressive (how do individuals who experience a state behave) or perceptive level (how do individuals, who we think are in a certain emotional state behave) - could provide a valuable insight to explore the implicit relationship between emotions and appraisals. Indeed, there are still uncertainties regarding this relationship. This is due, on one hand, to a not yet clear understanding of the degree of causality between emotion and appraisal and of the impact of intertwining appraisals. On the other hand, exploration is constrained by current methodologies. Researchers are limited in how they can capture the complexity of naturalistic emotional behaviour, which calls for a fine control of the experimental presentation of emotions/appraisals, and for the possibility of evaluating single nonverbal cues appearing along the global process of emotion formation.

Aims of the research and outline of the thesis

This PhD project focuses on the perception of expressions displayed by others. Meaningful recognition and interpretation of other individuals' internal states is essential in everyday interactions. The aim of this thesis is to provide a deeper understanding of the nonverbal cues that have an impact on the perception of emotions and associated cognitive appraisals. Concretely, this project explores the use of nonverbal cues by laypersons when these have to attribute mental states to others in interaction. The present research proposes to particularise facial emotion recognition in two ways.

First, we focus on the perception of human nonverbal behaviours. In a judgment study participants observe a person engaged in conversation with another one and have to attribute mental states, i.e. emotions and appraisals to that observed person. To tackle this problematic, a naturalistic video corpus (Lost Luggage; Scherer & Ceschi, 1997; 2000) has been selected. Videos present face to face interactions filmed unobtrusively: passengers claiming the loss of their luggage at an airport. For the present research, nonverbal cues are described with a technical coding system, FACS (Facial Action Coding System; Ekman, Friesen & Hager, 2002) that relies on the analysis of perceived facial muscle movements. This coding of facial movements is complemented by a coding scheme defined by ourselves to encompass the analysis of head, hand, arm and torso movements. These descriptions of behaviours and perceptive judgments allow us to make a link between mental state attributions and concrete physical expressions.

Second, the described behaviours are transposed onto a human-like virtual character. The use of a virtual character enables the fine tuning of facial expressions to test the impact of such expressions on the evaluation of the mental state of the observed person. In order to display appropriate facial and body expressions on the character, different elements of the character need to be adjusted, e.g., to manipulate the timing of onsets of different body actions independently from each other.

Expressions generated by the character are then judged by participants, who attribute mental states to perceived expressions using the same questionnaire as the one used for the human expression study. Particular cues of complex expressions are manipulated, one cue at a time. The impact of each cue's presence or absence is observed in terms of attributions. This clarifies to some extent the input of each element when placed in sequence of spontaneous behaviour. It also enables the creation of nuanced expressions for the character.

This two-step methodology is presented in the thesis as follows:

- In Chapter 2, *Theoretical Introduction*, we review the principal theories of emotion generation and expression, with a special focus on the componential appraisal theory. Specifically, we define the concept of appraisal as well as its relationship with emotion. Next, we discuss the importance of non-verbal cues and, in particular, facial and body movements in the perception of emotions and appraisal by third parties.

- In Chapter 3, *Appraisal components of emotion*, we explore the link between emotions and appraisals based on attributions by laypeople. We report the results of an empirical study in which participants were asked to watch videos of two persons interacting with each other, and assess the mental state of one of them through the attribution of labels which we measured on likert scales. We focus on emotions of joy, anger, relief, sadness, contempt fear and shame. A model of the co-presence between appraisals and these emotions is found through correlation and regressions. We compare our findings with the expected links predicted by appraisal theorists.

- In Chapter 4, *Expressions of appraisal and emotion*, we explore how emotions and appraisals are *perceived*. We use the same empirical study results as in Chapter 3, and moreover annotate the facial and body expressions present in each video. In this way, we can focus on the contingency between different nonverbal cues of interlocutors and the emotions and appraisal labels give by laypeople involved in the study. We analyse the correlations between nonverbal cues, emotions and appraisals, and compare our results with similar studies in the scientific literature.

- In Chapter 5, *Manipulated expressions of appraisal and emotion*, we reach the second part of the thesis, i.e., we perform a sensitivity analysis of the impact of single non-verbal cues to the perception of emotions/appraisals. To this end, we use a virtual character, Greta (Bevacqua, Prepin, Niewiadomski, de Sevin & Pelachaud, 2010), as a tool to finely tune the presence of single nonverbal cues shown to individuals. To enable the character to manage such precise coordination between behaviours as in complex naturalistic scenes, we define a representation scheme that encompasses the dynamics of emotional displays and we called it multimodal sequential expressions (MSE) language. MSE ensures the description in a formal way of the configuration of signals as well as of the relations that occur between them.

  We carry out a second subjective perceptive study, in which the participants are asked to evaluate the mental state of the character in a similar way as in the test presented in Chapter 3. The results of this second study, which we report herein, enable us to run some statistical analysis to look for differences between conditions with and without some particular facial cues in emotion/appraisals attributions.

Contributions

The main contribution of this thesis is to explore how emotions and appraisals are perceived through a two-step methodology, namely: emotion attribution by laypeople to **video recordings** of non acted behaviour in a face to face interaction; and **fine-grained analysis of** nonverbal cues through the use of a **virtual character**.

Specifically, our contributions include:

- *Use of spontaneous behaviour filmed unobtrusively in a real and naturalistic situation for the study of appraisal and emotion attribution by laypeople.* While a majority of the previous and ongoing perceptive studies are based on actor portrayals, on static stimuli or on experimentally guided emotional expressions, the videos used in the thesis are of a particularly emotional situation: the loss of a luggage at an airport. Passengers claiming the loss of their luggage were filmed by a hidden camera (Lost Luggage Case; Scherer & Ceschi, 1997). The use of such a naturalistic corpus enables us to verify predictions by presenting expressions of emotions in their complexity.

- *Use of a virtual character as a novel tool* for exploring spontaneous and non-stereotypical behaviour. Virtual characters and other virtual representations of humans have been used for perceptive studies on a large span of communicational issues (Niewiadomski & Pelachaud, 2010; Grizard, Paleari, Lisetti, 2006; Kang, Gratch, 2010; Kleinsmith, De Silva, Bianchi-Berthouze, 2006). Indeed, it has been shown that they can convey emotional states well (e.g. see Hyniewska, Niewiadomski, Mancini & Pelachaud, 2010) and have been used in different psychological studies (see Schillbach et al., 2006). Such virtual characters are an ideal tool as specific behavioural displays can be strictly controlled, with the possibility to manipulate cues one by one while all other cues are kept constant. Such precise manipulation of cues is not possible with human recordings, whether of actors or of real life situations. Nevertheless, none of the existing studies using virtual characters so far has enabled fine tuning of expressions by the experiment designer. The character used for this thesis enables manipulations in terms of FACS units for the face and body and torso mouvements.

- Development *of a model of* Multimodal Sequential Expressions (MSE) *turning the virtual character to a novel tool that can be used for studying* sequences of superposed face and body actions. For the purpose of this research, the character was modified to have the ability to generate actions with independent timings of onsets and offsets, whereas the majority of current characters rely on simple one attack-hold-delay for all actions that have to start and end at the same time during an expression. Our MSE model was based on the idea that emotional behaviours are ordered in time following specific rules. To define the rules we based ourselves on annotation of several corpora with very strong intensity expressions of emotions. Displays generated following our model were evaluated through perceptive studies that confirmed that such expressions are well recognised. Our studies have also shown that our MSE displays are better recognised than emotions presented through a static form, with expressions at their apex, and also recognised better than dynamic expressions not respecting sequentiality. Thus, we can say that our new tool for generating face and body movements enables us to go beyond animations displaying schematic and oversimplified expressions.

- *Analysis of the impact of single nonverbal behaviours in conveying emotions by means of fine regulation of facial expressions in a virtual character.* Many ongoing studies support the claim that it is the pattern of different components which contributes to trigger and express an emotion. In the thesis, we concentrate instead on some *individual* behavioural units. Without diminishing the importance of sequence, timing and interaction of behavioural cues, we reinforce through experimental evidence the impact of single facial cues. Through the use of a virtual character we show that removing one single facial cue from a complex expression leads to an overall change in the holistic perception of emotion and appraisal. This allows us to enrich the hypotheses of expression-emotion/appraisal contingency formulated in Chapter 3 through a correlation analysis, with a deeper causal connotation.

While concentrating on the major contributions stated above, this thesis additionally advances the current state of the art in the following respects:

- *Creation of an ad hoc body description scheme for the lost luggage dataset.* Given that no standard body movement coding system is widely accepted for emotion studies, a body action coding scheme has been created specifically for this data collection on "Lost Luggage corpus". The latter corpus was annotated with the formulated description scheme. Having an accurate description for the body movements enables us to make precise hypothesis on the relation between expressions and appraisals and emotions, which can be then verified in the active part of the study through the use of the virtual character.

- *Annotation of spontaneous facial expressions on muscle level* by a certified FACS coder. To our knowledge, this is the first video corpus acquired in a naturalistic emotion triggering situation in an everyday setting with a complete FACS annotation. The annotated corpus is available on demand to other researchers in the field who want to use it for their research.

1. Theoretical introduction

...les sens, l'imagination et la pensée elle-même, si élevée, si abstraite qu'on la suppose, ne peuvent s'exercer sans éveiller un sentiment corrélatif, et que ce sentiment se traduit directement, sympathiquement, symboliquement ou métaphoriquement, dans toutes les sphères des organes extérieurs, qui la racontent tous, suivant leur mode d'action propre, comme si chacun d'eux avait été directement affecté.

Gratiolet, 1865

1.1. Review of Emotion Theories

The definition of emotion is still debated today in the scientific literature.

According to Stumpf, an emotion is an intentional state of evaluating *"a state of affairs*", (Stumpf, 1899, translated by Reisenzein &Schoenpflug, 1992). This intentionality is what differentiates emotions from sense-feelings, like bodily pain and pleasure or aesthetic enjoyment, which are more directly related to the senses, like sound and colours are. In normal circumstances emotion has an adaptive function - it is not a reaction inherent to a stimulus that leads to an emotion, but the goals and subjective perception of an individual.

In general, definitions of emotion can vary greatly. However many researchers will agree on seeing emotion as a "hypothetical construct" (Scherer, 2000), involving several components in interaction, and which cannot be observed directly as an entity. In the Component Process Model (CPM), five components are described, and each could be seen as having a different function (e.g. Scherer 2000; Sander, Grandjean & Scherer, 2005): a cognitive component responsible of appraising events, i.e. the process of assessing events; a peripheral efference component which is regulating system; a motivational component responsible of the preparation and direction of action; a motor expression component responsible of communicating intentions; and a subjective feeling component responsible of monitoring the internal state (see Table 1).

In the Component Process Model, emotion is defined as "*an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism*" (Sander et al., 2005).

In this thesis, we will concentrate only on the cognitive and expressive aspects of emotion.

An important question regards the level of interaction between **emotion and cognition** and the degree of dependence of the affect on cognitive evaluations. Definitions of emotion are thus diverse and are based on various theories: independence of emotions from cognition, cognition and arousal as components of emotion, physical changes as triggers of emotion, the thalamus essential in the differentiation between emotions, neuromotor programs defining emotions, dimensional perception of emotion, and in the end cognitive appraisal as necessary for emotion and hardly dissociable from it.

1.1.1. The role of cognition in definitions of emotion

Emotion as independent from cognition

According to Zajonc (1980), emotion is independent from cognition. These two systems are often processed in different brain regions: feeling and emotion are considered to occur in lower brain centres, whereas thinking in higher ones. Zajonc argues that as a consequence of this separation and independence, an affective reaction does not necessarily involve considerable prior cognitive operations (Zajonc, 1980). The emotional system is considered to treat the available affective information, and then to influence the behavioural reactions, independently from cognitive processing of information. Feelings and preferences, both of which Zajonc considers to be affects, are considered to precede cognitions. One can like something or be afraid of it before knowing what it is, as Zajonc puts it (Zajonc, 1980).

When one tries to recall or recognise a stimulus, the first element to emerge is the affective quality of the original object, person, name or anything else. The early affective reaction is "*gross and vague*"; however it can influence the ensuing cognitive processes and behaviours. He suggest that all perception might be containing some affect, as no item appears to us as neutral, but we associate an affective tone to everything. We do not read *an article*, but an *exciting article* or a *trivial article*. To Zajonc, cognitions influence emotions only in later phases, but the first emotional reaction is autonomous. Affective reactions are considered faster than cognitive processing, therefore affect is erroneously considered to be postcognitive (Zajonc, 1980).

As an example of his claim that emotions can occur in the absence of relevant cognitions, Zajonc demonstrated experimentally that participants can form preferences for stimuli to which they were subliminally exposed, yet their ability to identify those stimuli remains at chance level (Zajonc, 1980).

Moreover, Zajonc's idea that initial responses to affective stimuli are automatic and do not require awareness of those stimuli has been confirmed in numerous ways. The functional magnetic resonance imaging (fMRI) of the brain, for example, has allowed the study of activation of the amygdala during presentations of emotional faces, seen in a way preventing explicit knowledge of the stimuli by the perceiver (Whalen, Rauch, Etcoff, McInerney, Lee, Jenike, 1998a).

As a consequence of the results of this study, emotion and cognition are considered to interact through images and symbols. According to Zajonc images are recognised faster than words, or as he explains, they elicit affect faster than abstractions like words (Zajonc, 1980).

Another researcher, Lazarus (1984), refutes Zajonc's theory by arguing that emotion only plays a part in one, unique, system of information processing. Cognition is considered necessary and sufficient for the subjective perception of emotions, so emotions result from cognitive evaluations (Scherer, 2001).

Zajonc rejects Lazarus' broad definition of cognition, which does not permit to differentiate between *"cognition, perception and sensation"* (Zajonc, 1984).

Cognition and arousal as components of emotion

Whereas Lazarus argues that cognitive appraisal precedes emotion (1984), and Zajonc that emotion is independent from cognition and can exist without it (1980), Schachter and Singer (1962) claim that cognitive appraisal follows emotional arousal. Their "two-factor theory", also called the "cognition-arousal theory", has been the dominant theory of emotion in textbooks for more than twenty years (Schorr, 2001).

In their theory, Schachter and Singer assert that the mind and the body sorely contribute to the experience of emotion. Individuals have "semi-conscious assumptions" about what appropriate feelings should be in a specific situation. Those assumptions are acquired on personal experience and social constructs, and they direct feelings and have a critical impact on what one really feels in a particular situation (Carlson & Hatfield, 1992).

In addition to this cognitive evaluation, all emotions also have a second component: the intense physiological arousal. Schachter and Singer believe that arousal is an indispensable component of emotions (1962), at least as important as cognitive evaluation, the latter being needed to differentiate emotions. As Lazarus, Schachter and Singer consider that appraisals, emotional arousal and behaviour are interconnected. However, in this matter, Schachter and singer endorse a different order to the phenomena than Lazarus: first comes an unspecified arousal, then a quick appraisal of the situation is needed to put a label on the emotion.

Schachter and Singer put their theory to test (Schachter & Singer, 1962). Following the idea that individuals experience emotions only when they are physiologically aroused, the authors injected some participants with an arousing drug, epinephrine (adrenalin), which acts between three to five minutes after the injection. It produced palpitations, tremor, flushes and accelerated breathing, all natural symptoms of a variety of emotional states. Some other participants were injected with a placebo. All participants were told that they were injected with a vitamin supplement called "suproxin" supposed to influence vision. The authors manipulated the emotion inference by confronting their participants with emotional behaviour of a confederate. Half of the participants saw euphoric behaviour and half angry behaviour. Results show that when participants were not provided with information on the actual effect of the injection, they were more susceptible to attribute the arousal to the confederate's mood. The authors claimed that their results show that both cognitive and physiological factors contribute to emotion, under certain circumstances cognition followed physiological arousal; and moreover that individuals assess their emotional state, in part, by observing to what extent they are aroused physiologically (Schachter & Singer, 1962).

Marshall and Zimbardo (1979), for example, reproduced the euphoric condition, while including differentiated levels of epinephrine. Their results show that physiological changes not attributed to the injection did not increase the participants' positive mood, that is the bodily changes were not explained by participants as induced by euphoric behaviour of the confederate. What is more, higher epinephrine dosage groups showed a tendency to significantly increase the negativity of the reports (instead of increased attributions of positivity).
Already in 1983, Reisenzein reviewed the later studies and concluded there is no support for the claim that "arousal is a necessary condition for an emotional state", nor that emotional states can result from labelling of unexplained arousal.

Physical changes as triggers of emotions

Another controversial duality in the definition of emotion concerns the body-mind distinction. Can we define emotions as specific patterns of physiological changes or rather as an evaluation of those changes in a specific context? In other words, what defines emotion, body or mind? Which comes first – the experience of emotion or physiological arousal?

James suggests that "*bodily changes follow directly the PERCEPTION of the exciting fact, and that our feeling of the same changes as they occur IS the emotion"* (emphasis in the original, James, 1884). According to this theory, visceral and motor reactions to a stimulus precede the emotion felt; more exactly the emotional experience is based on physiological changes. It is also the great number of changes taking place during an emotional phase that makes it impossible to voluntarily reproduce the integral expression of any emotion. An expression can be achieved on the muscular level, but it will fail with the skin and heart changes, etc., which are an essential part of emotion (James, 1884).

James emphasised that both the visceral reactions, such as a churning stomach, and the overt bodily motor reactions, such as trembling or striking, are the basis of emotion (James, 1884). A physiologist, Lange, had a similar perception of emotion, although he emphasised the vascular changes, such as changes in blood pressure, as central to emotion (repeated in Carlson & Hatfield, 1992). Thus, this approach has been named the James-Lange theory.

It is this claim, asserting that emotion is the perception of bodily changes, which has lead to the search for specific physiological patterns that would be specific to particular emotions.
It is also this approach that lead to the foundation of the Facial Feedback Hypothesis, which states that facial movements can influence emotional experience (Tourangeau & Ellsworth, 1979). A large corpus of empirical data is available in literature on the eventual modulating function of different facial expressions, which might facilitate the subjective and autonomic components of emotion (e.g. for a review see Soussignan, 2004; Soussignan et al., 2012). Numerous studies report moderate but reliable effects of voluntary produced facial movements on self-reports and experience labelling. Thus a more positively valenced judgment was induced when participants reproduced facial movements of a smile, whether it was overt (e.g. Laird, 1974) or through a covert procedure like holding a pen between the lips in a way to create the same lip configuration as in a smile (Strack et al., 1988). More negative judgments were also reported when induced by frown poses (e.g. Kraut, 1982).
Moreover, studies also reported a reduced experience of some affects (pain, enjoyment, pride) following suppression of facial expressions and more intense emotional experience following exaggerated expressions (e.g. Gross & Levenson, 1997). However inhibiting expressions of anger, disgust or embarrassment did not decrease self-reports of these emotions (e.g. Gross & Levenson, 1997, Harris, 2001; Gross, 2002).
Soussignan, on the other hand, focused on the quality of facial configuration, that is the correspondence between the activated muscle configuration by participants and actual expression expected for an emotion, and its impact on self reports of positive emotions and autonomic responses (2002). He tested if smile display reproduction with the pen-holding technique (Strack et al., 1988) were modulated by the activation of the eye region, orbicularis oculi (AU 6 in the FACS coding system) which is associated with genuine amusement. Orbicularis oculi activation was induced by smile lip movements of greater intensity (AU 12 of intensity D), in other words participants reproduced the pen holding task with intense lip movements to induce cheeks raising with furrows in the eye region as strong smiles have been noted to induce more cheeks raising than moderate smiles. Some participants produced high intensity smiles without AU 6. Soussignan compared participants with same (high) intensity smiles with and without AU 6 and his results show that AU 6 increased the intensity of positive reports attributed to viewed pleasant and funny videos (Soussignan, 2002).
The James-Lange theory also states that there are neither specific brain circuits nor regions involved in the experience of emotion (Davidson, Putnam, Larson, 2000).

The role of the central nervous system in emotion elicitation

Another physiologist, Cannon, has questioned the idea that there are no brain regions specialised in emotion. Cannon argues that physiological changes are non-specific; viscera are relatively insensitive structures, and therefore cannot provide any subtle or complex kinds of information needed to differentiate diverse emotional states. According to Cannon, very different emotional states are associated with identical visceral changes and the physiological changes are too slow to account for the rapid emotional state changes (Davidson et al, 2000).

Due to the fact that the viscera is not a possible source of information for the emotion elicitation and differentiation, Cannon concluded the brain and more specifically the thalamus is the control centre of emotional behaviour. Therefore, according to Cannon, it is the activation of the thalamus that produces the emotional experience and the bodily changes associated with it.

In their experiments, Cannon and Bard have demonstrated that there are specific neural circuits involved in the expression and experience of emotion (Bard, 1928, 1929; cited in Davidson et al, 2000). Emotions seem to be closely tied to the brain organisation.

Neuromotor programs

When tackling emotions, other researchers define these as the result of "neuromotor programs", thus biologically fixed. These programs are believed to act independently from cognitive evaluations (e.g. Ekman, 1972). In 1962, Tomkins proposed a theory claiming a limited, fixed number of "basic emotions". Although each basic emotion can vary in intensity, they are considered discrete and based on a single "neuromotor program" whose triggering produces all the associated components of the specific emotion, such as vocal and facial expression, subjective experience of the emotion and physiological changes (Russell et al, 2003). According to some scholars, the great variety of emotional experience is due to the blending of the basic emotions (Hager & Ekman, 1983).

Basic emotion theorists consider emotions to be independent from cognitive processes. They also assume emotional expressions to be "truthful" information of the emotional state of the sending person, which are automatically interpreted by the "receiver" (any third party that can perceive signals or other cues emitted by a person). The decoding process is thus considered innate (Izard, 1994) and fast, therefore unconscious (Ekman, 1997).

The theory of basic emotions has been challenged by several theorists which lead to the development of new perceptions of emotion (Russell et al, 2003).

Dimensional perception of emotion attribution

Some theorists suggest that the internal state of a sender is defined as a point or region in a multidimensional space, along several broad bipolar dimensions (Wundt, 1874/1902/1904). Already Aristotle discussed emotions as varying gradually along one essential dimension that he defined as the "hedonic tone" (for a review on the hedonic dimensions see Labukt, 2012).

The notion that emotions may have several dimensions has already been described by Spinoza. According to this philosopher, emotions are not only varying in the dimension of valence (attainment), but also in their intensity (necessity) and arousal (persistency) (Spinoza, 1677/1955).

Nowadays, the majority of theorists argue that emotions are perceived in terms of two dimensions of valence and activation (e.g. Plutchik, 1980; Bradley & Lang, 1994, Russell et al, 2003).

Russell and colleagues conceptualise emotion as undifferentiated, differing on the dimensions "valence", that is the pleasure-displeasure dimension, and "activation", that describes the level of energy in a state or a person as sleepy-hyperactivated. According to this approach receivers agree with one another in the judgment of an emotional state on the two dimensions, whether they judge facial or vocal expressions (Russell et al, 2003). In fact, the focus is primarily on the subjective feeling, rather than on the underlying cognitive or physiological mechanisms and efferent expressions.

Pollick and his colleagues (Pollick, Paterson, Bruderlin & Sanford, 2001) who work on the perception of affects in body movements also explain data they observe through the use of a bi-dimensional space. Although the authors start by asking participants to attribute affect labels to point-light displays reproducing human arm movements, results lead them to the use of scaling algorithm used on confusions observed in the attributions. Scaling leads to a two-dimensional space, with a first dimension explaining 70 % of the variance (activation), and a second 17 % of the variance (valence). The psychological space of affects the authors obtain conforms to a circumplex structure with the arm movements scattered on a circle around the origin (Pollick et al., 2001).

Appraisal theories of emotion

The premise of appraisal theories of emotion is that cognition, or more precisely cognitive appraisal, is necessary for emotion and directly linked to it.

The term *appraisal* was first used by Magda Arnold (1960), although some researchers trace back the roots of this approach to Aristotle (see Reisenzein & Schoenpflug, 1992; Colombetti, Thomson, 2006). In his De Anima, Aristotle stresses the importance of integrating form and matter on all levels (Kafetsios & LaRock, 2005), seeing emotions as embodied cognitions (logoi en hylê; Aristotle, reprint 1981). In his Rhetoric, Aristotle (350 BC, reprint 1954) argues that emotions involve rationality, instead of being a hindrance in reasoning and appropriate behaviour.

Another influential thinker to advocate that emotions facilitate appropriate behavioural reactions is Kierkegaard (1849). He argues that emotions are what tie us to our personal goals and to surrounding reality. They help us relate to what exists, they are evaluative judgments of the world. To him, emotions are entirely thought-dependent.

Thus, we can define appraisal theories as approaches in which cognitive evaluation is at the core of emotions. As most recent appraisal theorists agree, these cognitive evaluations can be conscious or unconscious. Arnold was the first one to use the term "appraisal" to describe *direct, immediate and intuitive* evaluations as causal entities in emotion (Arnold, 1960). She argues that organisms keep monitoring the changes that can happen in their environment and judge whether these could be of relevance for their well-being. What is being evaluated is an entire environmental context of an organism, whether some beneficial or harmful stimuli are present in it, whether an ongoing event enhances or hinders the fulfilment of needs and goals of an organism (e.g. Arnold, 1960; Smith & Ellsworth, 1985).

Ortony and Turner (1990), Scherer (1992), and Kaiser and Scherer (1998) are authors who argue that there are a large number of highly differentiated emotions. Moreover, emotional states are not the effect of motor programs, but of appraisal processes. Thus, they counter the concept of basic emotions as fixed biological programs (for references, see Kaiser & Wehrle, 2001).

Most appraisal theories cite similar appraisals as crucial for the differentiation of emotions. Following Ellsworth and Scherer's (2003) comparative overview of major appraisal dimensions as postulated by major appraisal theorists, several central appraisal groups can be highlighted: novelty, intrinsic pleasantness, certainty, goal relevance, agency, coping potential and compatibility with norms (Frijda, 1986; Roseman, 1984; Scherer, 1984; Smith, Ellsworth, 1985; Ellsworth & Scherer, 2003). The majority of the theorists agree on the definition of particular appraisal groups (see Ellsworth & Scherer, 2003).

This research paper is focused on the appraisal theories, which are examined more in details hereafter.

1.2. Appraisal: Component process model of emotion

One of the new conceptions of emotion is based on the multicomponent evaluative process as it is adopted by the appraisal theories of emotion. These theories convey the idea that it is the significance that we attribute to different components of an event that creates an emotion. Arnold (Arnold, 1960) and Lazarus (Lazarus, 1960) pioneered in this psychological domain by saying that an emotion is elicited and differentiated by a subjective interpretation of events and situations.

It is not the situation or the stimulus as such that determines the emotion but the way an individual relates to the environment. This relationship is created through appraisal (for reference see Kaiser & Wehrle, 2001).

According to Scherer and his colleagues (1984), appraisal is made of a fixed sequence of "stimulus evaluation checks" ("SECs"). Their theory predicts intermediate expressions based on sequential appraisal checks and predicts an accumulation of patterns of expressions in the final expression.

The SECs are organised in terms of four appraisal objectives (Scherer et al, 2001):

1. how relevant is this event to the person, how directly does it affect the person (relevance)
2. what are the consequences of this event, how does it influence the person's well-being, the immediate and long-term goals (implications)
3. how well can the person cope or adjust to these consequences (coping potential)
4. what is the significance of this event with respect to the person's self- concept and to social norms and values (normative significance)

1.2.1. Relevance

Given that it is essential to process any new piece of information that could affect the organism's well being and one's immediate and long term goals, organisms keep monitoring their internal and external environment.

One of the first appraisals of the environment checks how **sudden** is the occurrence of an event, how rapidly it was triggered. The suddenness of a stimulus requires from an organism a fast allocation of resources into the processing of the change that occurred. Given the limited capacities of cognitive processing that can take place at a given time, attention is selective. What is more, some stimuli require more in-depth processing than others (Desimone & Duncan, 1995). Suddenness is such a criterion that leads to a shift of attention towards the involved stimulus (Yantis, 1998). The appraisal of an event judged as sudden is highly expected to be linked with fear (Scherer, e.g. 1999), which is the appraisal with the fewest predicted links to emotions (Scherer, 1999).

Organisms also evaluate the likeliness of the occurrence of an event, based on general knowledge and on past experience. They check if stimuli relevant for their well-being are consistent with their expectations in terms of frequency of occurrence given a specific context and of the time of occurrence. Thus organisms scan the environment for stimuli that are **relevant and incongruent** with their expectations. Any unpredicted event or any absence of a predicted one requires organism's attention to figure out the eventual consequences and calculate the associated risks (Ellsworth & Scherer, 2003). Frijda has put a particular emphasis on relevance by stating that before all *emotions are relevance detectors* (Frijda, 1986).

1.2.2. Implications

The appraisal of **goal obstruction** (e.g. Scherer, 1993) is considered (under various names) as a major component of fear, anger and sadness and linked negatively to joy (Scherer, 1999). It was named goal hindrance (Scherer, 1999), motivational discrepancy (Smith, Lazarus, 1993); motive inconsistency (Roseman, 1984), undesirable event (Ortony et al., 1988) and goal blocking (Izard, 1977). Goal obstruction happens when the satisfaction of a need is put out of reach, delayed or the amount of effort has to be increased (Ellsworth & Scherer, 2003).

1.2.3. Coping potential

Organisms have to face events and, in case of relevant changes in their environment, they take active steps to diminish the impact of occurring obstacles. Hence, it is important for an organism to evaluate its capacity to **cope** with ongoing events. This includes the evaluation of control that the organism has on the event (the extent to which it can influence the event, e.g. have an impact on the cause) and of the organism's power to change such event.

Having some coping potential, or in other words having some control on what is happening, is another evaluation that might be considered an important component of anger according to some theorists (Scherer, 1993; Ellsworth, Smith, 1988; Lerner, Keltner, 2001). Some studies have shown nevertheless that the link between coping potential (at least the perceived aspect of it) and anger is not so univocal (Frijda et al., 1989; Roseman et al. 1990).

Coping potential is expected to be very low in fear and somewhat low in sadness (Scherer, 1999).

1.2.4. Normative significance

Two appraisals with a social dimension have also been put forward, one on internal and one on external norms. For humans, which are a social species, respecting others is important. Shared rules, or so called norms, are essential for the good functioning and well-being in everyday situations (Ellsworth & Scherer, 2003). Perceived injustice, in other words the **violation of external standards** or immorality is expected to be linked with anger, when others transgress the norms, and guilt when it is the individual who transgresses it (Scherer, 1999).

The second social appraisal, the **respect of internal standards** or self-consistency (Scherer, 1999) is an evaluation of an individual's behaviour in terms of one's self-ideal or internalised moral standard (Ellsworth & Scherer, 2003). Low respect of internal standards is expected to be related strongly to anger and very strongly to shame and guilt (Scherer, 1999).

In Table 1 hereunder, different studies have been combined to contribute to the definition of the link between particular appraisals and emotions. We report how several emotions are differentiated in terms of their appraisal components.

Table 1

Emotion Differentiation Based on Stimulus Evaluation Checks

	Emotions						
Appraisal Components	Anger	Sadness	Contempt	Fear	Joy	Relief	Shame
Suddenness	Open	Open	Open	High	Open	-	Open
Goal Obstruction	High	High	Open	High	Very low	Low	Open
Relevance/Discrepancy	High	Open	Open[b]	High	Open	Low[a]	Open
Coping Potential	High	Low	High	Very low	Medium	-	Open
Violation of External Standards	High	Open	High	Open	Open	Open	Open
Respect of internal standards	Low	Open	Open	Open	Open	Open	Very low

Note. "Open" signifies that different evaluation results are compatible with the respective emotion and/or further detail (on subtype of emotion and/or subchecks) needed for predictions.

Joy, Anger, Sadness, Fear and Shame are based on Table 5 (Scherer, 1988).

[a] *Discrepancy for Relief is based on Table 2 (Scherer 1988).*
[b] *Discrepancy outcome for Contempt is based on Table 1 (Scherer, 2001).*

In this model, emotion is defined as an episode of interrelated, synchronised changes in the state of all or most of all of five components, as a result of the four categories appraisal checks.

As appraisal is a process, checks are evaluated more than once. The sequence theory postulates that the information of a prior check is necessary for the following checks to take place. Thus, organisms constantly have to observe their environment, to evaluate and re-evaluate the changes, running through the SECs sequence again and again, in a fixed order, as described above.

The outcomes of all the SECs are always subjective and depend exclusively on the appraising individual's perception of an event and its inferences.

According to Scherer (e.g. 1987; 1988), emotion is considered to consist of five components, or subsystems corresponding to five distinctive functions and five organismic subsystems, as shown in Table 2 below:

Table 2

Organismic Subsystems and Related Components of Emotion

Function	Subsystem	Component
Evaluation of Stimulation	Information Processing	Cognitive
System Regulation	Support	Neurophysiological
Preparation and Direction of Action	Executive	Motivational
Communication of Reaction and Intent	Action	Expressive
Monitoring, Attention Focusing, and Reflection	Monitor	Subjective Feeling

Note. Table from Scherer (1987)

Contrary to the basic emotion theorists' assumption that emotions are recognised straightforwardly, appraisal theorists expect an interaction between numerous aspects of an expression, whether it is vocal or facial. According to this multicomponential approach, the outcome of each check is seen to affect all four different emotion components: action tendencies, autonomic processes, motor expression and subjective feeling (Sander, Grandjean, Scherer, 2005). These four components are in constant interaction during the process of generation and of differentiation of emotions, as the organism constantly evaluates and re-evaluates the surrounding situation on the basis of these checks. Thus, emotional state itself is never static, but in constant change.

In natural settings, interactions between individuals are emotionally coloured. In humans, the whole organism keeps on disclosing the affective reactions to the environment, ranging from low level cues, as the blinking rate or respiration changes (e.g. Gomez, Shafy, Danuser, 2009) or the rate of vibration of vocal folds (Bachorowski & Owren, 1995), to motor expressions (Sato & Yoshikawa, 2007). Since the 1960s (Plutchik, 1962; Tomkis, 1962), emphasis has been directed towards the face, to which a major communicative role is attributed (Mehrabian, 1971), starting with Tomkins and McCarter's work (1964), up until today (e.g. Kaiser & Wehrle, 2001). Attention has also been more and more directed towards vocal parameters expressing emotions (Banse & Scherer, 1996; Brosch, Grandjean, Sander & Scherer, 2009; Schuller et al., 2009).

Nowadays, however, the scope of interest has shifted towards emotional changes occurring at different levels: the multimodal encoding and decoding of emotions. This topic of emotion expressivity is covered in the next section.

1.3. Emotion expressivity

The field of emotion expressivity focuses on the triggering and expression process of emotions. In the first part of this section, we describe how researchers define emotion expressivity. Emotion expressivity has been investigated on the level of nonverbal behaviour: face, voice, body posture and movement, multimodal channels. Researchers have also investigated on how individuals decode expressions they see in others, and on how information from the face, eye gaze, bodily behaviour and posture can be decoded.

In the second part of this section, we describe some findings on the processing of emotional nonverbal behaviours, whether on conscious or unconscious level.

1.3.1. Theoretical approaches to emotion expressivity

According to Scherer, emotional states are "*almost always accompanied by a motor expression component*" (Scherer & Ellgring, 2007). The readability of the body actions and poses of a person enables the observer to infer the internal states and attitudes of such person (Scherer & Ellgring, 2007). Among different emotion theories, two important approaches have been proposed, that diverge in their respective understanding of emotions triggering and expression. Both propose explicit predictions for emotion-specific facial expressions. While one conceives emotions as categorical, the other describes them as componential entities.

- Discrete emotion theories (Ekman, Friesen, 1975; Tomkins et al., 1964, Tomkins, 1982; Ekman, 1972) focus on a small number of so-called basic emotions, in particular anger, fear, joy, disgust, sadness, happiness, shame and guilt. These are considered to result from innate neuromotor programs and to produce a fixed behavioural response. This expressive response is unitary in nature, emotion-specific and universally recognised.

- Componential appraisal theories to emotion (Scherer & Ellgring, 2007; Sander, D., Grandjean, Kaiser & Wehrle, Scherer, 2007; Roseman, Smith, 2001; Turner, Ortony, 1992; Scherer & Ellgring, 2007) on the other hand, stipulate that the individual elements of facial expressions are determined by appraisals of a given situation.

Discrete emotion theories

According to the discrete emotion theoreticians, emotions are triggered by automatic mechanisms, such as neuromotor affect programs. Studies following this approach focus on a few prototypical patterns. Tomkins (1982), for instance, described these affect programs as leading to some expressive patterns specific to particular emotions. The expression of emotions was considered, from the ancient times to the nineteenth century, to be universal across ages and cultures (Russell, Bachorowski, Fernandez-Dolls, 2003). The number of these "basic emotions" is limited. As regards facial expressions, discrete (basic) emotion theory states that they are direct displays of internal states and that the ability to decode them in term of basic emotions is innate (1994a) and quick, thus considered unconscious (Ekman, 1972).

So far, basic emotion theorists have not published concrete predictions for facial expression of prototypical emotions. However an "Emotional dictionary", which is part of FACSAID manual created for the interpretation of FACS (Facial Action Coding System; Ekman, Friesen & Hager, 2002), has been developed by Ekman and colleagues. It lists possible meanings and links between FACS Action Units and Emotions. This dictionary is accessible but remains unpublished.

Componential emotion theories

Cognitivist theoreticians following the componential approach to emotion counter the concept of discrete emotions resulting from automatic and biologically fixed programmes (Kaiser & Wehrle, 2001; Kaiser, Sherer, 1998; Ortony, Turner, 1990). They advance that the variability and complexity of emotion expression and experience can be understood without any reference to basic emotions. According to these theoreticians, there is a great number of very differentiated emotional states that are captured by the labels only through a process of grouping of different states, through some kind of averaging and central tendencies. Scherer names these "averaged" states "modal emotions" (see for example Scherer & Ellgring, 2007). The appraisal theory's predictions for these modal emotions are the same as the ones suggested for basic emotions.

The specific element of the componential appraisal approach is that emotions are considered to be consequences of cognitive evaluations (appraisals). An emotional state would result from the significance given to different elements of an event. Consequently, an emotion is not defined and triggered directly by a situation or a stimulus, but depends from the relation established between a person and surrounding environment. This relation is created through appraisal (for a review see Kaiser & Wehrle, 2001).

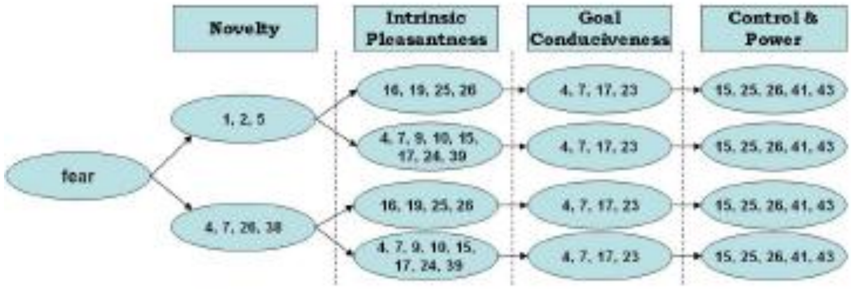


Figure 1 ***Different AU that could express emotion of Fear, according to appraisal theory (adapted from Scherer, 1987, cited by Paleari, Grizard, Lisetti, 2007)***

Although this process relies on a succession of evaluations, some authors emphasise that this cognitive evaluation can happen in an automatic, fast and non-conscious manner (see Sander, Grandjean, Kaiser, Wehrle & Scherer, 2007;).

1.3.2. Nonverbal expressions of internal states

Internal states, such as emotions, have been studied in terms of triggers as well as in terms of efferent effects on motivational and motor tendencies. Bell was a precursor with his writing on the relation between states of the mind and expressive body behaviours (1844). According to some researchers, expressions can be seen as "*rudiments of adaptive behaviour, which have acquired important signalling characteristics*" (Kaiser & Wehrle, 2001). Thus, expressive behaviour is a socially influenced message, prone to be regulated by the organism, as well as a true externalisation of internal states.

Emotion has been investigated in relation to different expressive modalities. The major weight of research has focused on the perception of emotion as displayed on human face (Kaiser & Wehrle, 2001; Kaiser, Scherer, 1998) and this since the nineteenth century, with pioneering work by Duchenne (1876) and Darwin (1872/1998). However numerous research underline the importance of decoding and encoding through other modalities. Some scientists investigated the voice (e.g. Scherer, Banse, Wallbott & Goldbeck, 1991; Grandjean et al., 2005; Schuller, 2002; Belin, Fillion-Bilodeau & Gosselin, 2008) others- body movements (Wallbott, 1998; Pollick, Paterson, Bruderlin & Sanford, 2001).

Inner states and the human face

Facial behaviour decoding

Duchenne (1876) was the first to investigate ways in which individual muscles contribute to perceived facial changes. Through electrical stimulations of specific facial muscles, *localised electrisation* as he called it, he turned researchers' attention, including Darwin's, towards *how* emotions are communicated through facial expression.

Nowadays there is no doubt left on the fact that the face is a privileged place for the expression and the decoding of emotions (see Scherer, 1992). Facial expressions are often the cue element in efficient communication and in general a rich source of information, considered sometimes even to be the modality transmitting the greatest amount of information (Mehrabian, 1971). Facial expressions have also been explored in the search of basic emotion patterns (Ekman, 1972; Ekman, Friesen, 1976).

Facial expression processing activates specific brain regions, which are independent from processing of stable features of the face. The processing of the latter has been shown to involve specialised brain regions within the ventral occipito-temporal cortex, i.e., the fusiform gyrus (Allison, Puce, Spencer,, McCarthy, 1999; Kanwisher, McDermott,, Chun, 1997; Vuilleumier, Armony, Driver, Dolan, 2003) with the fusiform face area (FFA) being specifically activated by presentation of faces, independently from presentations of any other category of objects, such as body cues, symbols, etc. Other regions involved are within the lateral occipital and temporal cortices (e.g. Kanwisher et al., 1997).

Whereas the location where the face processing occurs is known, the mechanisms underlying facial expression processing are still debated, and so is their location. Some studies showed that facial expressions are closely linked to emotional processing in general, even if each of the emotions seems to be processed independently from the others (Goren & Wilson, 2006). As any emotional stimulus, facial expression is also believed to be processed by the amygdala. According to Sato and his colleagues (2004), amygdala activation is furthermore modulated by the interaction between different facial emotional expression and face direction.

Today, Fridlund (1994) as well as Russel and Fernandez-Dols formulate the question whether facial expressions should be considered as "direct expressions" of underlying emotional processes, that is "readouts", or rather as social signals.

One model, which follows Wundt's work (1902) and Bühler's (1968), considers facial expressions to have both functions (Kaiser & Wehrle, 2001). According to this approach, facial expressions are the residues of adaptative behaviours, which have acquired signalling properties. Facial expressions are thus considered reliable external manifestations of internal states and social signals used in interpersonal communication.

Eye gaze behaviour decoding

An element closely linked to facial expressions is the direction of **gaze**. The gaze direction is also considered to be a crucial social signal, as it indicates which object a specific individual is evaluating and where the focus of such individual's attention is. It also regulates turn-taking in face to face interactions, enables to regulate intimacy and to exercise social control (Kleinke, 1986). Nevertheless, behaviour of eyes is not sufficient to transmit a social message, such as threat or friendliness – it has to be interpreted along other, contextual, cues (Adams, Kleck, 2003).

Perception of eye gaze is fundamental in the evaluation of where the other person's attention is (Langton, Watt, Bruce, 2000); this is an important contribution to the feeling of personal implication of a perceiver, in the inferences of other person's intentions (Baron-Cohen, 1995; Langton et al, 2000) and in the processing of a face. Behavioural data suggest that the direction of eye gaze is critical in different face processing tasks such as categorization. For example, gender categorization of (non emotional) faces is faster when the eyes have a direct gaze than when they have an averted one (Hood, 2003). Furthermore, other person's gaze direction is also important in the detection of threat in the environment (Baron-Cohen, 1995; Adams et al, 2003; George, Driver, Dolan, 2001) as in the case of a gaze averted from the perceiver or direct as in the case of threatening social contact.

Many studies suggest that the brain regions mostly involved in the processing of eye gaze are the superior temporal sulcus and adjacent areas of the temporal lobe and the amygdala (George et al, 2001; Holmes et al, 2006).

Expression and gaze decoding

Facial behaviour and eye gaze have also been studied in the affective sciences domain in terms of how the interaction of the two modalities influences the perception of emotion by third parties. Several studies have shown an interaction between emotional expression observed in a third person and gaze direction of that person. studies confirms higher interaction between emotional expressions and direct gaze (Sato et al., 2004), other find that interaction is more substantial in case of an indirect gaze (e.g. Adams et al., 2003).

Sato and his colleagues (2004) confirm the effect of interaction between the emotional expression and the eye gaze direction of a perceived face. Angry expressions directed toward viewing participants elicited more negative emotional experience than the angry gazes averted from the subjects, as measured on a likert scale as the strength of the emotion that subjects felt when perceiving the stimuli models' expression. The emotion recognition, defined as the strength of the emotion that subjects recognised from the stimuli models' expression, however, showed no significant interaction. Thus, in an angry face with a direct gaze the perceiver was more engaged in the emotion, but did not consider the expresser to have stronger anger, than an angry expresser with an averted gaze.

The same study claims that the amygdala is sensitive to this interaction, with a greater activation for angry expressions looking toward the perceiver than those looking away, but it is not the case for neutral faces, as depicted using fMRI (Sato et al., 2004).

Sato interprets his results claiming that the amygdala is involved in the processing of the significance and the relevance of the facial expression to the perceiver, but is not related to the processing of physical features of the face (Sato et al., 2004).

Adams and his colleagues' results (2003), on the other hand, report a different pattern of activation. The authors claim that gaze direction differentially modulates the perceptual clarity of emotional facial expressions. The authors find greater amygdala activity for angry expressions with averted gaze than angry with a direct gaze. They explain that the amygdala is responsible for threat detection and for the detection of the source of threat, particularly when it is ambiguous. This explains why angry faces with a direct gaze and fearful faces with an averted gaze are more quickly and accurately recognised than angry with an averted gaze or fearful with a direct gaze, the latter being considered an ambiguous threat. However, according to the authors, it is the ambiguity that increases amygdalian activation.

Inner states and the human body

Studies show that emotions can be recognised by other humans from body movements (De Meijer, 1989; Wallbott & Scherer, 1986; Dael et al., 2012) as well as from posture (Coulson, 2004). Some researchers also stress the importance of multimodal expressions (Scherer & Ellgring, 2007).

Bodily expression of emotion

There is evidence that body movements and static body postures can transmit emotional information with high levels of accuracy (see for example early work by Walters & Walk, 1986). However, compared to the input of voice and face information on other's emotional state, research on the bodily expressions has been surprising less predominant in scientific literature on affects. It was unrightly unsufficiently attended to, given clear evidence from psychotherapeutic settings (see for example early work on body congruence and psychotherapeutic rapport by Scheflen, 1964, and work on the perception of psychotherapist-client rapport by laypersons based on observed bodily expression indices described by Trout and Rosenfeld, 1980).

Sometimes, before the 90's, body positions were claimed to carry little information about specific states (Ekman & Friesen, 1974). Claiming that emotions are not differentiated by their body movements alone, Ekman and Friesen (1969a) acknowledged only that information about tension and in consequence about expressions' intensity is provided by gesticulation, but also and more importantly by body alignment and position.

Later, Wallbott (1998) claimed that rich information on emotion is transmitted through behaviours' expressivity, which consists of movements' visible characteristics like fluidity or speed. He also claims that this behaviour's expressivity plays an important role in communication, i.e. on the decoding of information like emotion, mood or personality, which it undeniably facilitates.

Wallbott (1998) claims that there exists a relation between quality and quantity of the state (mental, emotional and/or physical), which are inseparable from the intensity. Moreover, in his work he shows that behaviour like gesticulation gives information not only on what (e.g. through the hand shape), but also on how it is being communicated (e.g. through the motion quality features, such as amplitude or speed of the behaviour).

Several researchers ran perceptive studies on categorised motion features (Wallbott & Scherer, 1986; Gallaher, 1992; Johansson, 1973; Pollick, 2004). Wallbott and Scherer (1986) asked twelve students to judge actor portrayals of joy, anger, sadness and surprise using three point likert scales (low-1 to high-3) defining movement characterists (fast, expansive, energetic, active, pleasant). Interrater agreement was high. Significant differences were obtained for "fast", "energetic" and "pleasant" scales between emotions. This result was mostly due to the difference between sadness ratings and the three other emotions. In sadness portrayals actors used less energetic and less active mouvements. In sadness, mouvements were also less expansive.

Sadness was characterised by greater head orientation down or away from the partner and more frequent hand mouvements. The majority of present hand mouvements were shrugs and self-manipulators.

Gallaher (1992) asked participants to judge peers behaviours in everyday situtions on 78 items defining characteristics of how movements are performed as "lethargic", "jerky", "graceful", etc. A factor analysis was performed on these items and four dimensions were retained: expressiveness, animation, expansiveness, and coordination. The four dimensions were internally consistent, stable over time, and stable across raters (Gallaher, 1992).

Pollick and his colleagues (Pollick et al., 2001) observed how participants are able to interpret light-point displays reproducing human movements in terms of affect labels. One set of studies of affect perception involves judging point-light displays of arm movements performing knocking and drinking actions. Results show confusions between several emotion labels and the authors applied a multidimensional scaling procedure (Kruskal, Wish, 1978) to explain affect label attribution patterns. The scaling algorithm used on confusions lead to a two-dimensional space, with a first dimension explaining 70 % of the variance, and a second 17 % of the variance. The authors identified the first dimension to correspond to activation and the second to valence. They suggest that affect in body movements can be seen in terms of activation related to movement kinematics while valence of the movement is carried *in the phase relations between the different limb segments* (Pollick et al., 2001).

Coulson (2004) focused on joint rotations and studied how six of them (head/neck bend, chest bend, abdomen twist, shoulder forward/backward, shoulder swing, and elbow bend) could help recognising six emotions (anger, fear, happiness, sadness, surprise, and disgust). Students judged computer generated avatars and the author identified which variables could contribute to the attribution of particular emotional states. While disgust postures were recognised by less than 50% of participants, other emotional expressions showed high agreement rates. Coulson characterised each state in terms of static posture features. Happiness, for example, is characterised by a head backward position and no forward movement of the chest. Arms are raised above shoulder level and are straight at the elbow.

Kleinsmith and Berthouze (2003, 2006, 2007) looked at posture in the affective computing setting with one of their aims being endowing systems with the ability to recognise body postures of their users. In their study, participants judged acted postures of four emotions (happiness, sadness, anger and fear) with high recognition rates. The authors created an afective posture recognition system that maps the set of postural descriptors into affective categories (2003). Their results show a correct automatic categorisation of acted angry, happy and sad postures while relying only on static descriptions.
In 2011, Kleinsmith, Berthouze and Steed extended their previous study to non-acted behaviours, collected in the context of a Nintendo Wii video game. Japanese students who judged these static behaviour displays on faceless avatars and had to attribute a label describing the state of the player. Judgments showed above chance level agreements. The labels used were grouped: concentrating (determined, focused, interested); defeated (defeated, having given up, sad); frustrated (angry, frustrated); and triumphant (confident, excited, motivated, happy, victorious).
The important low-level posture description features for emotion state differentiation are mainly the arms and the upper body (Kleinsmith et al., 2011). Based on participant attributions, results show that forward and backward movement of the torso differentiated between the more “active” states (frustrated and triumphant) and the less “active” states (concentrating and defeated). The body is slightly bent forward in the concentrating and defeated postures, whereas it remains upright or slightly bent backward in the frustrated and triumphant postures. The state “frustrated” was the only one not to be recognised above 50% and according to the authors it may require kinematic information in order to be recognised by external observers. Defeated and Triumphant expressions were very well differentiated based on three main features: vertical and lateral amount of opening of the body and bending of the arms (Kleinsmith et al., 2011).

Some studies also show that interactions between modalities have an impact on how affective displays are perceived. Social psychology studies (Kimble and Olszewski, 1980; for a review see Sato, Yoshikawa, Kochiyama & Matsumura, 2004) have demonstrated that the direction of gaze modulates the significance of an emotional message. Other studies show also a modulation of the perception of emotional messages by an interaction between facial and head cues (Hess, Adams & Kleck, 2007; Krumhuber, Manstead & Kappas 2007; Sato et al., 2004) and facial and body cues (e.g. Aviezer et al., 2008, Van den Stock, Grèzes & de Gelder, 2008).

Bodily behaviour and posture

There is some evidence that third parties interpret body behaviour in terms of affective states (e.g. Walters & Walk, 1986; Coulson, 2004). Although already mentioned by Darwin (1872/1998) to encode different qualities of emotions, with emotion-specific patterns, body behaviour was claimed by some researchers (Ekman & Friesen, 1974) to transmit nothing more than the quantity or intensity of emotion. This channel has been neglected in comparison to the face.

Body movements can be encapsulated by the term **kinesics,** which was firstly used by Birdwhistell (1952), a pioneer in the study of nonverbal behaviour**.** In kinesics, we differentiate between action behaviours and positions. Action behaviours are supported by the positioning of the body. They are *discrete units of body actions which are not part of body positioning and which have relatively easily discernible onsets and offsets* (Harrigan, 2005). Body positions are basic units defined as the alignment of one or a set of articulations (Dael et al. 2012). They change relatively infrequently and generally less frequently than body actions (Harrigan, 2005).

In kinesics, the majority of attention has been focused on action behaviours rather than body positioning. The studied actions refer mostly to the gestures related to the regions with the greatest amount of movement frequency, the head and the hands (Harrigan, 2005).

Former coding systems of body movement

Since the 1950's, some observational coding systems have been developed based on anatomical features (through a segmentation of the body into parts based on the skeletal system), directional features (spatial dimension of movement), movement quality (including the quality of dynamic movement or body shape), and/or body positions and actions units. As Dael formulates it (2012), "*observational coding is based on explicit operational definitions of an a priori defined set of behaviour codes and follows a fixed coding procedure"*. The idea is to provide a protocol precise enough so as to limit the subjective judgments and enable replicability of the coding by different coders, while not being too detailed or too time consuming.

The different coding systems could be used for annotation performed manually by humans or automatically computed.

A major reference in the domain is Labanotation (Laban, 1956, 1966, 1975) that Laban himself called Kinetographie (Laban, 1928). Although initially created exclusively for the annotation of movement in dance, it is today a standardised system for the analysis of human movement. Its ambition is to analyse every aspect of motion as precisely as possible. Motion is recorded in a form similar to the musical notation system, where actions of the different body parts are recorded in time. Symbols are used to analyse quantitative and qualitative features of movement, such as direction, duration and intensity.

Birdwhistell (1952, 1970), on the other hand, designed a notation system based on linguistic principles. He felt that gesture is structured in units similar to those in language. Analogous to phones, allophones, phonemes, and morphemes, he proposed kinesic counterparts termed kines, allokines, kinemes and kinemorphemes.

Coulson (2004) created a framework for describing spontaneous expressions of emotions through posture. He treats the body as "a system of interconnected rigid segments, roughly corresponding to the bones connecting the major joints". He describes relationships between these segments in terms of rotations about one or more axes of the joints connecting them. Although, as the author underlines, there are fifteen major joints in the human skeleton with a total of twenty-nine degrees of freedom (ankles, knees, elbows and the chest are monoaxial, wrists are biaxial, and shoulders, hips and the head/neck and abdomen 'joints' are triaxial), it is interesting to simplify the model to make it more manageable by reducing the number of postures. Coulson opts, in his coding system, for thirteen segments and nine degrees of freedom. The upper body consisted of seven segments (head/neck, chest, abdomen, two shoulders/upper arms and two forearms), the lower of six (two thighs, two shins and two feet). The degrees of freedom relate to axes of joint rotation. Notwithstanding the many ways body's center of mass can move, Coulson opts for the encoding of three levels: forward, backward and neutral, as he claims that for emotional states the body is not required to indicate anything more than moving towards or away from a relevant stimulus.

Kleinsmith and Berthouze (2003) suggested some groundwork for a FACS-like formal system to code posture. They proposed a general description of posture based on angles and distances between body joints that would be particularly useful in the recognition of affective postures. Basing themselves on Laban's 'sphere of movement' (Laban, 1966), the authors use *eighteen* kinematic features to describe postures. Their description relies on of limb-to-torso distances and overall expansion of the body in the frontal, lateral, and vertical dimensions. Nine joints were considered: head, neck, collar, shoulders, elbows, wrists, torso, hips, and knees.

Kleinsmith, Berthouze and Steed (2011) associated a vector containing a low-level description to postures. Their description is built upon 3-D joint Euler rotations recorded by the motion capture system. Each rotation value was normalised to [0, 1] by taking into account the fact that the maximum range of rotation differs for each joint and most joints can't achieve 360° rotations. For some of the joints, the range of one direction of the movement is greater than the range for the opposite direction of the movement, for example for the forward and backward mouvement of the hip.

Recently, a coding system called BAP (Body Action and Posture) was proposed by Dael and colleagues (2012) for the study of emotion expression. Particular attention was paid to the discrimination between body postures and actions. The BAP coding system provides a fine-grained description of body postures and actions along three main directions: sagittal, vertical and lateral. Body postures were coded both at a local level (focusing on one body segment such as the left arm or the head) and at a global level (considering the posture of the whole body) with a particular focus on the upper body parts. Functional classification of hand gestures was also introduced in their coding system.

Body movement coding systems are applied in different fields of research such as nonverbal communication and interaction, dance, as well as for the expression of emotion through the body. Most of the time, the coding is not organised theoretically or conceptually in a way to allow one to deal in an intuitive way with a different context that the one aimed for by the authors (Harrigan, 2005). For the particular context of emotion expression, no single standardised and widely agreed upon way of coding body movements existed at the beginning of this PhD work. At present, the BAP system, as well as the system proposed by Kleinsmith and Berthouze are on their way to being tested in different emotional contexts and hopefully lead or contribute to an accepted system such as FACS is for the face.

1.3.3. Perception of emotional expression by third parties

Conscious and unconscious emotional processing

When a third party observes behaviours of a person, the emotional message these behaviours eventually convey could be perceived on a conscious or on an unconscious level. Consciousness may depend to some degree on attention attributed to stimuli. In general, emotionally loaded stimuli create an attentional bias – emotional stimuli being of importance to the organism, they attract more attention than the neutral stimuli so as to enable further processing (e.g. Williams, Mathews & MacLeod, 1996; Mogg & Bradley, 1999).

Anderson (2005) has studied the attentional prerequisites for consciousness of the stimuli with the attentional blink paradigm. The attentional blink is the impairment in report of the second target presented briefly after the identification of first stimulus. Anderson demonstrated that emotional stimuli presented second diminish the attentional blink, and it is the arousal value of the target that modulates the degree of visual awareness. This finding is an important argument for the automaticity of emotional processing.

Qualitative or quantitative difference

There is a debate growing in magnitude about whether conscious and unconscious processing present a qualitative or a quantitative difference (Phillips et al., 2004). Phillips and colleagues have demonstrated in their functional magnetic resonance imagery (fMRI) study a substantively different pattern of activation for the two states of awareness. In overt, conscious facial presentations of fear and disgust appears a double dissociation: fear increases the activation of the amygdala, but not of the insula, whereas disgust activates the insula, but not the amygdala. On the other hand, in unconscious, covert, facial presentations of fear and disgust, the authors note no activation of the amygdala for fear and no activation of the insula for disgust. These findings support the idea of distinct neural correlates of conscious and unconscious emotion perception.

Nonetheless, some authors do find an activation of the amygdala for covert fear against happy faces (Whalen, 1998). The 33ms presentations of emotional stimuli (167ms of neutral face masks), with 8 of 10 subjects reporting they had not seen the facial expressions, putting in doubt the above reasoning as a proof of a differentiated processing pattern.

Critchley and his colleagues (Critchley et al., 2000) have also investigated whether different pathways are activated during explicit ("effortful") and implicit processing of facial emotional expressions. The implicit processing was defined differently than in the above experiments, with the cognitive awareness being manipulated (reduced in the implicit condition) and not the visual awareness as in back masking. During a gender categorisation task participants were not attending explicitly to the emotional expressions. Using fMRI the authors have confirmed dissociation in the neural substrates, with implicit processing of fear and happiness activating the amygdala region, and their explicit processing activating the temporal lobe cortex.

Automaticity and the amygdala

Emotions are essential for the generation of adaptive responses to the environment. As the environment provides more stimuli than our brain can process, significant emotional stimuli, such as potential threats, have to be evaluated fast, sufficiently and appropriately. A commonly held view is that emotional mechanisms are automatic (Zajonc, 1980; Whalen, 1998) at least to some extent. In a reflexive manner, the allocation of attention is greater for emotional stimuli, as illustrated by the emotional Stroop task, where the naming of the colour of a word is slowed down for emotional words (Williams et al., 1996). This interference is "involuntary", hence somewhat automatic. Similarly, a stimulus loaded with emotions, such as a spider among flowers, is detected faster in a visual search task than non-emotional stimuli (Öhman, Flykt & Esteves, 2001).

Automatic information processing is probably mediated by a subcortical brain pathway, involving the amygdala (Öhman, 2005; Adolphs, 2004). This region is an almond-shaped nuclear complex located in the limbic system in the anterior part of the medial temporal lobe. It has widespread connections to many other brain regions. These nuclei are heterogeneous and do not constitute an integrated functional system.

The amygdala was called a "relevance detector" by Sander, Grafman and Zalla (2003). According to the authors it is not only crucial for the processing of threat related information, but also all kinds of other biologically relevant stimuli, including the processing of relevant social cues.

The amygdala is crucial for providing direct and indirect top-down signals on sensory pathways, which can influence the representation of emotional events, especially when related to threat. Threat related matter needs to be treated preferentially compared to other incoming information. The solution is for the processing to take place in parallel and without awareness, that is automatically (Vuilleumier, 2005; also see Shafer et al., 2012).

Accordingly, research on the importance of social interactions brings us to face evidence that social-related cues tend to dominate in different tasks where processing resources are limited (see Adolphs, 2003). Face understanding and recognition for example are critical for the identification of others and therefore their processing is of great importance as the face has many features to be examined. Most researchers agree that eyes are a privileged spot of attention (Yarbus, 1967; Whalen et al., 2004; Taylor, George & Ducorps, 2001; Itier, Alain, Sedore & McIntosh, 2007) and that the direction of gaze plays a key role in the interpretation of a facial expression.

1.4. Studying emotional behaviours: choice of audio-visual corpora

In the study of emotion expression, the great majority of studies rely on videos of actors (see Bänziger & Scherer, 2007; Gosselin, Kirouac & Doré, 1995).

Another approach is to collect experimentally guided behaviour. One way is to involve participants in a slightly passive situation, such as emotion eliciting films (e.g. Hess, Banse, Kappas, 1995) or presentation of emotion inducing stimuli, e.g. Belfast Spaghetti data (see Douglas-Cowie et al., 2011). To increase implication, some more researchers have begun to use interactive computer games to elicit emotions by manipulating appraisals (e.g. Kappas & Pecchinenda, 1999; MacDowell & Mandler, 1989; Kaiser & Wehrle, 2001; Kaiser, Wehrle & Schmidt, 1998). This method takes advantage of the possibility to systematically manipulate the rules and situations the players are confronted with, thus eliciting sets of appraisals thought to lead to different emotions. Audio-visual recordings of such emotional interactions have been used in studies by Kaiser and Wehrle that analysed facial reactions of participants to particular appraisal-defined situations (e.g., Kaiser & Wehrle, 2001).

In the assessment of emotion and appraisal processes through questionnaires, imagery tasks have been in use, e.g. by recall of past events. Participants relive in their imagination past emotional experiences and respond to questions about their evaluation of these experiences (e.g. Ellsworth & Smith, 1988; Scherer, 1993; Smith & Ellsworth, 1985). Following that approach, a narrative emotion-sharing task was created and recorded by With and Kaiser (With & Kaiser, 2011) for the study of nonverbal and verbal emotional behaviours.
The method entails the advantage that it can easily be employed for a large number of persons in a controlled experimental setting. Nevertheless, imagery tasks appear not to be the best way to study the expressions of inner states and how these states are organised. It has been criticized for their questionable ecological validity and their overreliance upon memory. It has been claimed that retrospection about past affective situations leads to more cognitive judgments than the assessment of an undergoing affective situation (e.g., Fredrickson, 2000).
Moreover, according to Parkinson and Manstead (1993), self-reports obtained with these methods reflect prototypical socially shared knowledge on the content of different emotions rather than actually occurring emotion-eliciting processes (see also Frijda, 1993).

Another approach suggested as an alternative for actor or experimental studies is to rely on live television shows. Although debatable (see Scherer, 2003), it has been used in some emotion expression studies (Douglas-Cowie et al., 2003).

The observation of naturally occurring emotions seems to be the most ecologically valid solution. Although the assessment of cognitive and emotional processes by questionnaires or interviews, during or directly after the emotional episode (e.g. Folkman & Lazarus, 1985; Scherer & Ceschi, 1997; Smith & Ellsworth, 1987) has been adopted by researchers, the recording of behaviours occurring in such emotional situations is scarce. While such situations provide extremely valuable information and this method has the advantage to be ecologically the most valid, recording has rarely been applied because of evident practical problems. It is rather difficult to observe and assess emotional episodes that naturally occur for a large number of persons.

Following that last approach, the video corpus we use comes from a field study in which individuals in an emotional situation have been recorded. The Lost Luggage corpus was recorded at Geneva's international airport by Scherer and Ceschi (Scherer & Ceschi, 1997; Scherer & Ceschi, 2000).

A few other field studies of spontaneous behaviour in emotional situations have also been realised and discussed in literature. Bonanno and his colleagues (2002), for example, looked at childhood sexual abuse and disclosure of traumatic events. Kraut and Johnstone (1979) recorded on paper (looking through binoculars) or through pictures taken through a photo lens the smiling behaviours of bowlers and hockey players, respectively.

2. Problematics 1: Appraisal components of emotion

About, my brain!

Shakespeare, 1602

2.1. Introduction: Cognition as part of emotion

Although all appraisal theorists agree on the importance of cognition in emotions, with predictions concerning which appraisals contribute to which emotions, the nature of the association between the two phenomena is still debated.

Appraisals can be considered either as preceding emotions or as components of emotions. Is it the evaluation of a situation that is essential in triggering an emotion or is the evaluation already part of the emotion? Or to go even further, is appraisal already the emotion itself? Is it the definition of an experienced emotion?

Parkinson (1997) distinguishes two levels on which the nature of the relation between appraisals and emotions has to be explored. First, he suggests examining whether the association is conceptual or empirical, and second he proposes analysing whether the association is that of **contingency**, **necessity** or **sufficiency** of appraisals for emotion's presence.

To answer the first questioning, appraisal theories argue that "actual real time processes are empirically associated in some way with actual real-time emotional processes" (Parkinson, 1997). However, the link that is observed might come from definitional issues (Scherer, 2000). Testing empirically whether appraisals are a cause or a component of emotion may not be conclusive simply because this perception could be relying on individuals' perception of the link. Appraisals may not be causal, but may be perceived as such by individuals facing their everyday events and may contribute to the recognition of emotional states in oneself and in others. Parkinson (1997) stresses that studies claiming associative or causal connections between appraisal and emotion were relying on self-report methods, which focused on the convergence of attribution of both concepts. Hence, Parkinson (1997) argues that the link between appraisal and emotion is to be placed rather at the perceptive than the definition level. He claims that appraisals can be seen as signals on which individuals base themselves to attribute emotions, that appraisals "may be viewed as conveying the communicative content of emotions".

As regards Parkinson's second analytical level, whether on the conceptual or empirical level, the association between appraisals and emotions can be defined as having three possible natures: contingency, necessity or sufficiency (Parkinson, 1997).

First, an appraisal can be seen as simply **contingent** to emotion. Thus, appraisal could be associated, only on occasions or only partly, with the phenomenon or the concept of emotion or with emotion attribution. When differentiating between different emotions, one can say that *sometimes* specific emotions could be characterised by their different patterns of appraisals. If we postulate that the relation between appraisal and emotion is only of contingency, appraisal is to be seen as only one of the factors contributing to the definition or perception of emotion (e.g. Parkinson & Manstead, 1992).

Given that some observations show that emotions are not characterised by any necessary defining components (Fehr & Russell, 1984; Russell & Fehr, 1994), but by loose clusters centred around "prototypical" exemplars (Rosch, 1978; Scherer, 1994), some theorists claim that the connection cannot be anything more than contingent. In fact, the majority of appraisal theorists share the idea that appraisal and emotion are linked on the empirical level and argue mostly that *evaluative aspects of emotional experience constitute a definite stage in an information-processing sequence leading to the affective reaction itself* (see Parkinson, 1997). They assume that a specific appraisal is not sufficient for an emotion to be considered present, but that an emotion is determined by the combination of different components present at the same time without claiming that one particular appraisal is necessary for the presence of particular emotions (Kuppens et al., 2003).

The second stance is to see an appraisal as **necessary** for an emotion to be present or to be attributed. In such case, an emotion is always associated with an appraisal and on the particular level, different emotions are always associated with distinctive appraisal patterns (see Ortony, Clore & Collins, 1988; Reisenzein, 1994). Some appraisal theorists claim such an invariant one-to-one relationship between some appraisal combinations and certain specific emotions (Roseman & Smith, 2001; Smith & Pope, 1992). According to Roseman (2001), for example, sadness is the result of appraisal of the situation as inconsistent with goal enhancing (appetitive) expectations, high certainty, low coping potential and the event being attributed to some impersonal agency. Besides others, Roseman (1991) used vignettes manipulating the content of appraisals described to see the impact on emotion attribution and saw that appraisal interactions had highly significant effects on emotion ratings.

The third stance is to consider appraisals as necessary and **sufficient**. Each emotion is completely characterised (or determined) by appraisal and different emotions are entirely characterised by different appraisal profiles. Such connection implies, to a certain extent, an overlap of the two processes of appraisal and emotion. It is interesting to note that such a characteristic type of appraisal, that is sufficient to produce an emotion and that in fact always does produce an emotion, would seem to carry some affective information in itself already. Lazarus is one of those to assert such a very direct and causal link between appraisal and emotion (Lazarus, 1991). According to him, evaluative judgment is the only determinant of emotional experience so his point of view can be placed on the third level determining the nature of association.

2.2. Research question

In this *Naturalistic expression study*, the relation between appraisal and emotion attribution by third parties is examined.

In order to define how an appraisal is connected to an emotion, we could analyse its necessity and sufficiency. Intuitively, we deem an appraisal as *necessary* for an emotion to occur if the emotion can be present *only if* the appraisal is present as well. Conversely, an appraisal will be *sufficient* to produce an emotion if its presence implies the manifestation of the emotion.

We have selected six appraisals (suddenness, goal obstruction, coping potential, relevance and importance, respect of internal standards, violations of external standards) and seven emotions (Joy, Anger, Relief, Sadness, Contempt, Fear and Shame) to be attributed. Appraisals were chosen for their expressive value described in literature, i.e. Scherer and Ellgring give a clear prediction in terms of action units for each of these appraisals checked as present. Besides, appraisals were matched to emotions in which they are considered to be present (Scherer & Ellgring, 2007). Relief has been added to counter the bias of a unique positive emotion.

Our appraisal-emotion exploration is made of two parts. First, we look at the nature of the relationship between emotion and appraisal that are linked together by theory, by exploring if it is of necessity, sufficiency or simply of contingency. Second, we explore the link on the contingency level and we look at linear relationships between attributed appraisals and specific emotions, basing our expectations on appraisal literature.

Thus, to look at the nature of the relation between appraisals and emotions, we follow Kuppens and his colleagues' (2003) approach that focused on the specificity, necessity and sufficiency of different appraisals contributing mainly to anger. Kuppens and his colleagues (2003) relied on recall of laypersons' personal experiences, with instructions asking for recall of emotional episodes characterised by different appraisal profiles, some of them being present or absent in recalled events. The authors show that none of the four appraisals that they call goal obstacle, unfairness, control and other accountability, are necessary or sufficient for the experience of anger.

With regards to appraisal-emotion associations, three types of connections have been proposed; namely: conceptual, descriptive and causal (Parkinson, 1997). In this thesis, we adopt the descriptive point of view by describing the *empirical* association between appraisal and emotional processes. To illustrate this perspective, we could see appraisal processes as terms of a dictionary, and emotions as sentences. Necessity implies that sentences (emotions) have to be formed with the terms of the dictionary (appraisals), while sufficiency means that some appraisal words are enough to form a proper sentence/emotion. Notice that in the first case, the dictionary is not necessarily unique, i.e. an emotion/sentence could be formed with the aid of one or more additional dictionaries. Therefore, emotions need an appraisal-based description, *and possibly* further elements (e.g. physical elements).
An interesting case is when the connection is a necessary and sufficient one (equivalence). In our example, the sentence/emotion can be entirely written with only the words of the dictionary/set of appraisals. This suggests that emotions can be described completely as a set of appraisal processes (possibly, appraisals happening with a certain order). Notice that this is however weaker than equivalence at the conceptual level (emotions *are* but a set of appraisals) or at the causal level (appraisals completely determine emotions and emotions are completely determined by appraisals).

We also explore the connection between the two concepts of emotion and appraisal, yet we base our analysis on the representation of emotion and on the attribution of mental states by laypersons to third parties. Thus, we look at the specificity, necessity and sufficiency of appraisals for emotions in the context of perceived behaviours by third parties. We want to see how attributions by third parties (based on judgment of video clips) follow the rules described in scientific literature, based on reported experienced emotions (self-reports of emotion).

The link between appraisal and emotions is also explored through simple correlations and regressions. We will compare our results with the CPM predictions (see Table 1 in Review of Emotion Theories).

2.3. Predictions

Following the theoretical background reported in section 1.1. Review of emotion theories and summarised in Table 1 (*Emotion Differentiation Based on Stimulus Evaluation Checks, p. 23) we advance the following hypotheses.*

2.3.1. Theoretical hypotheses

Based on the previous review of the literature, we hypothesise that there exists a relationship between attributed appraisal labels and attributed emotion labels, as apparent in third party ratings of spontaneous behaviour in short video clips.

2.3.2. Operational hypotheses

On the basis of the previously discussed Component Process Model by Scherer, we expect no appraisal to be strictly sufficient for the attribution of an associated emotion.
We expect no appraisal to be strictly necessary for the attribution of an associated emotion.
Based on the summary presented in Table 1, we expect correlations presented in Table 3.

Table 3

Prediction of Correlations Based on Table 1

	Emotions						
Appraisal Components	Anger[1]	Sadness	Contempt	Fear	Joy	Relief	Shame
suddenness	0	0	0	+	0	n.a.	0
goal obstruction	+	+	0	+	-	0	0
relevance/discrepancy	+	0	0	+	0	-	0
coping potential	+	-	0	-	0	n.a.	0
violation of external standards	+	0	+	0	0	0	0
respect of internal standards	-	0	0	0	0	0	-

Note. "+/-" signify a positive/negative correlation expected, "0" signifies no correlation expected.

[1] Emotion labels are capitalised to to mark the difference between third party attributions and affective phenomena experienced by individuals and to visually differentiate them from appraisal labels.

2.4. Analyses

To observe if an appraisal is sufficient or necessary for the attribution of an associated emotion, necessity and sufficiency indices were calculated respectively for all appraisals and emotions.
To observe what kind of contingencies appear in the attributions, correlations were run between appraisals and emotions.
To observe which appraisal attributions seem to contribute most to emotion attributions, regression analyses were performed.

2.4.1. Necessity and sufficiency of appraisal indices

We examined the relation between appraisals and emotions in terms of necessity and sufficiency. First, to follow the approach used by Van Mechelen, Smits and De Boeck (2003) and Gara and Rosenberg (1979), attribution scores of appraisals and emotions were dichotomised. For emotion scores, attributions of 0 were kept as 0 and all other turned into 1 (0 = no emotion, 7 = strong emotion). Given that the appraisal scale was bipolar (0 = totally did not agree with the statement on appraisal, 7 = totally agreed on the statement), scores ≤3 were considered as showing the absence of an appraisal and therefore coded as 0, and ≥4 was seen as showing presence of an appraisal and coded as 1.

After dichotomisation, instead of quantifying the strength of the association by one standard phi (φ) coefficient for the correlation between the presence-absence of appraisal and emotion, two indices were calculated for each of these two variables: necessity and sufficiency. Following Kuppens, Van Mechelen, Smits and De Boeck (2003; Gara & Rosenberg, 1979) we defined this degree more formally in terms of a φ index, which we calculate for necessity and for sufficiency.

We define the degree of necessity as $\varphi_{\text{necessity}} = \dfrac{P(\text{appraisal}|\text{emotion}) - P(\text{appraisal})}{1 - P(\text{appraisal})}$

In this equation, the denominator ensures that $\varphi_{\text{necessity}} \leq 1$. If the conditional probability $P(\text{appraisal}|\text{emotion}) = 1$ then $\varphi_{\text{necessity}}=1$. This corresponds to stating that for a given pair (appraisal, emotion), whenever the emotion is present, the appraisal will be present as well. In other terms, that appraisal is *necessary* to form the emotion.

On the other hand, $\varphi_{\text{necessity}} = 0$ if $P(\text{appraisal}|\text{emotion}) = P(\text{appraisal})$, which means that the appraisal is independent from the emotion. In other words, when $\varphi_{\text{necessity}} = 0$ an emotion can occur whether the component under examination is present or not, i.e. that component will be *unnecessary* to form the emotion. Other values of $\varphi_{\text{necessity}}$ measure to which extent the component is needed to form an emotion. A value can be either positive or negative. A positive index implies that the presence of an emotion increases the chance of appearance of an appraisal. Conversely, a negative index implies that the apparition of emotion decreases the chance of appearance of an appraisal.

Similarly to the necessity equation, we define the degree of sufficiency as:
$$\varphi_{\text{sufficiency}} = \frac{P(\text{emotion}|\text{appraisal}) - P(\text{emotion})}{1 - P(\text{emotion})}.$$

This second equation is interpreted the same way as the first equation of necessity. Thus, $\varphi_{\text{sufficiency}} = 1$ implies that the occurrence of a certain appraisal is sufficient to produce a given emotion.

2.4.2. Correlations

Spearman's rho (ρ) coefficient (bivariate correlations) are calculated for the appraisals scales and emotions scales. Correlations were run for the predicted relations (one-tailed) and also to explore non predicted ones (two-tailed).

2.4.3. Regressions

According to appraisal theories (e.g. Scherer, 1988) each emotion is characterized by a specific appraisal profile. This postulate implies that emotion attributions of our participants should be differentiated on the basis of their responses to the appraisal attributions.
To test this hypothesis we conducted Standard multiple linear regression analysis.
The scores values to the appraisal attributions are the predictor variables.

2.5. Materials

As will be remembered from section 1.4, a particularity of the presented approach to the emotion-appraisal problematic lies in the evaluation of audio-visually recorded behaviour occurring in a natural setting.

In our study we use a situation close to everyday situations and we evaluate how laypeople judge the mental state of observed individuals.

Relying on actor portrayals or experimentally collected data could lead to biased results through the accentuation of stereotypical behaviour. Our data relies on unobtrusive recordings from a hidden camera and show face to face interactions. By using naturalistic data we try to obtain more subtle expressions. We are also prepared to see compound mental states, more complex emotions, which we want to measure through non-exclusive likert scales (all likert scales are independent).

2.5.1. Stimuli Creation

The chosen corpus presents an emotional situation, as it shows passengers claiming their loss of luggage at an airport (Scherer & Ceschi, 1997; Scherer & Ceschi, 2000). The Lost Luggage corpus presents face-to-face dyadic interactions: one camera was filming unobtrusively the passengers and one was aimed at the hostess. The images from the two cameras have been synchronised in order to be shown together.

In the Lost Luggage corpus 112 passengers were filmed, of which 40 were selected mostly based on the quality of the collected video material. In the original field study (Scherer & Ceschi, 1997; Scherer & Ceschi, 2000) only the smiles were observed: smiles with raised cheek bones (AU 6+12) and the smiles alone (AU 12 without AU 6). The films in the corpus focus on the passenger, with a head and torso framing, while showing in the right corner a reduced size video of the face of the hostess.

As we wanted to evaluate the attribution of appraisals and emotions to one mental state at a time, clips had to be cut in order to present one mental state per segment. The first issue to be resolved was to do a segmentation of emotional extracts. We wanted to find, through a somewhat non-subjective method, when an emotional state starts and when it ends. The original corpus included one minute long video clips (16-bit colours), that have not been cut in a way to depict only one mental state per segment.

Judges have been asked to annotate videos and to signal in time all mental states and state changes. The task was explained through guidelines that were provided in a written format that was additionally read orally to make sure the participants thought about all the provided examples, without dropping lines.

Judges were asked to view a video clip and to annotate by means of Anvil, a software enabling time-aligned audio-visual annotation.

They were told that their task was to indicate changes between different mental states of one person (see Appendix I). They had to select a period of time (by indicating a starting and an ending time) for each state and to define this mental state. To avoid guiding participants into a particular theoretical framework, guidelines provided examples of action tendencies, motivational changes, appraisal attributions and emotional labels. Participants were told orally that the focus is on “internal states” of passengers that have lost their luggage and that the films come from a hidden camera at an airport. They were told that in one video clip a passenger can display several mental states and moments of neutrality and that they had to indicate them all. They could describe what they see in sentences, through expressions or labels either orally (transcribed by the experimentator) or in a written format on a piece of paper or directly in the provided ANVIL software, with which they were assisted.

Several such tools have emerged in the last few years that are important for human behaviour annotation, as they allow defining a representation scheme adapted to the particular corpus/data to be processed. For all tasks of annotation we opted for Anvil (ANVIL, Video Annotation Research Tool. http://www.dfki.de/~kipp/anvil/), sustaining the use of parallel, time-aligned, tracks for the coding of different modalities following numerous criteria.

Seven laypeople, administrative stuff from the technical university, have been invited to be judges for the task. The two first judges that to participate (an account officer and a junior secretary) have thought that the task was extremely difficult. They gave the following reasons:

- it is impossible to say that a state is changing;
- it could be possible to point that there is an emotion such as anger in one moment, but not point to a time;
- these individuals are talking and not experiencing and even less expressing some emotions; and
- saying what the passengers feel or in what state they are, without being guided by specific emotional labels is difficult.

Consequently this procedure has been dropped for laypeople (the evaluations of five remaining judges that were invited to take part in our study were not recorded) and instead four experienced judges have been invited into the evaluation. By experienced judges we mean individuals who have developed some acuity in the perception of facial expressions due in three cases to their professional activity (virtual character synthesis; facial graphics; FACS coding) and one recruited for his interest in the nonverbal communication in general. All the four understood the task straight away from reading the guidelines.

Each clip was annotated by three experienced judges.

In case of ambiguity, for example when one expert out of three considered less changes in a clip than the other experts, and made a segment last longer, we opted for leaving out the non agreed upon segment. To reformulate, the solution was, when possible, to cut the clip in a "more restrictive way", by eliminating moments that lead to discordance. Only moments on which judges agreed to display only one state were kept as one piece. If a state starting during a movement or a sentence was preceded by a neutral phase, a second or a second and a half might have been added to the chosen segment to enable the display of the movement development.

In two cases in which ambiguity did not allow an easy and straightforward cutting even in the above, restrictive, manner, a fourth experienced judge was asked to annotate the video clips. In both cases, two judges annotated long segments and one judge a much shorter segment. The fourth judge had a very similar segmentation to the short segmentation, for the two concerned videos. Thus we followed this restrictive segmentation, as it enabled a definition of mental states to be extracted and presented in separate clips.

After cutting, 64 clips were obtained. Several extracts from these were excluded from the corpus, as they involved a fragment where the face was majorly obstructed or hidden behind glasses that reflected light in the view of the camera, or were presenting a situation outside of the original canvas (e.g. talking to a third person). The final choice consisted 41 clips, of 4 to 56 seconds duration, with a majority of 20-28 seconds. These were encoded with a temporal resolution up to 1/25th of a second.

2.5.2. Online study creation

The *Naturalistic expression study* was created using an open source software, lime survey v. 1.71+ (http://www.limesurvey.org/), that enables to generate online questionnaires. The software works with a php/MySQL database and requires a local or server installation. The study can be accessed through a simple web browser, however to protect the database the access was blocked to be accessed only internally. It was presented to the students on a GNU/Linux (Slackware) system. The clip-stimuli were presented on full screen (in the original size) on LCD 19 inch screens (1280 x 1024 pixel resolution). Participants used headphones.

Given that it was important to have participants concentrating on the task and not taking their gaze away from the screen during the display of stimuli (especially as some extracts were less than 10 seconds long), participants were encouraged to stop whenever they felt tired or less motivated. For that same reason we wanted to reduce the number of presented stimuli so as not to lead to a mental overload and dissipation. A pre test was run on three non-naive participants (two women and one man collaborators) with 6 video clips and the participants reported the amount to be of an appropriate length to maintain, not straining their concentration.

Consequently, we created 7 rating blocks of six stimuli clips (in three semi-randomised orders).

2.6. Sample

A pre test was run on 57 students (47 men) from a technical university (Institut Universitaire de Technologie de Montreuil). No student came from psychology or affiliated sciences, the majority came from computer science. Women were a minority in this technical school and showed a non-homogenous population (great age distribution, differences in domain of studies, and reported psychologist/psychiatrist history, such as anxiety and depression or intake of substances).

In consequence, for the main study, men participants were encouraged to take part and they were the only ones kept for later analyses.

122 students from the technical university took part in the main study. Only male participants were kept for analyses (N = 98).

2.7. Procedure

2.7.1. Set-up

The study was run in a computer room and participants arrived in groups from two to ten. Each participant accessed the study individually through a web browser. Participants were spread out in the room. The guidelines provided on the first web page were sufficient for the understanding of the tasks. Participants viewed short videos and answered the same set of questions after each video.

2.7.2. Questionnaire presentation

Each evaluation set consisted of 16 judgments: nine questions and seven attributions of emotion intensities (see Appendix II).

On the first page after the video display, questions evaluate suddenness, goal obstruction, detection of an important and incongruent event, coping potential, respect of internal standards and external standards violation. Participants answered on a 7-point likert scale, ranging from 0 = totally disagree to 6 = totally agree. With such a measure, an average significantly above 3 can be considered to confirm the presence of an appraisal and an average significantly below to confirm the absence of an appraisal in a given clip. Appraisals were presented as cited above, in the chronological order defined by the Componential theory (e.g. Scherer, 2001).

The participants also had to judge, on the second page after the video, whether the observed passenger is experiencing Joy, Anger, Relief, Sadness, Contempt, Fear and Shame. Each emotion is evaluated by participants on a separate 7-point likert scale ranging from zero (no emotion) to six (strong emotion). An emotion can be considered attributed when the average score is significantly >0.

Participants were randomly attributed to rating blocks. Emotional labels were presented in two controlled orders, the same order of presentation being kept for all stimuli judged by the same participant.

Participants watched and evaluated from 6 to 42 short video clip extracts, depending on their concentration level and their willingness to participate.

2.8. Study design

In our *Naturalistic expression study* we recorded two classes of dependent variables (DV):

- scores measured using 7 likert scales of emotions: Joy, Anger, Relief, Sadness, Contempt, Fear and Shame
- scores measured using 6 likert scales of appraisal: suddenness, goal obstruction, important and incongruent event, coping potential, respect of internal standards, violation of external standards.

For both DV scores range is between 0 and 6 points.

The independent variable (IV) is the audio-visual corpus comprising 41 video extracts presented to participants.

2.9. Results

Participants

98 students from a technical university took part in the study. Each participant evaluated between 6 and 24 video clips. Each video clip was evaluated by 11 to 70 men. They were aged between 17 to 25 years (m = 19, SD = 1.51) and the great majority spent their last five years in France.

Dependent Variable (DV)

Appraisal scales have a distribution from zero (rejection of a statement) to 7 points (total agreement with a statement).

Emotion scales have a distribution from zero (no emotion) to six (strong emotion).

Table 4

Distribution Values for Appraisal Scores

Distribution Values	Appraisal Components: suddenness	obstruction	discrepancy	coping	respect internal standards	external standard violation
Mean (m)	4.27	4.44	4.21	3.64	3.78	4.46
95 % Confident Interval for Mean						
Lower Bound	4.16	4.34	4.10	3.54	3.68	4.37
Upper Bound	4.38	4.55	4.31	3.75	3.88	4.56
Standard Deviation	1.971	1.880	1.889	1.904	1.787	1.714
Minimum	0	0	0	0	0	0
Maximum	7	7	7	7	7	7
Skewness	-.559	-.601	-.483	.079	-.101	-.710
Kurtosis	-.764	-.646	-.682	-1.107	-.830	.049

Note. A score of 0 indicates a non-attribution of an appraisal (the negation of a statement) and 7 indicates a high attribution of an appraisal (strong agreement with a statement).

All clips taken together, the lowest appraisal attribution mean is for coping potential (m = 3.64; SD = 1.90) and the highest is for external standard violation (m = 4.46; SD = 1.71).

Table 5

Distribution Values for Emotion Scores

	Emotions						
Distribution Values	Anger	Sadness	Contempt	Fear	Joy	Relief	Shame
Mean (m)	1.47	1.43	1.35	1.61	.39	.56	.91
95 % Confidence Interval for Mean							
Lower Bound	1.38	1.33	1.26	1.51	.33	.49	.83
Upper Bound	1.56	1.52	1.45	1.70	.45	.63	.98
Standard Deviation	1.594	1.665	1.717	1.699	1.120	1.202	1.371
Minimum	0	0	0	0	0	0	0
Maximum	6	6	7	7	7	7	7
Skewness	.989	1.134	1.213	.883	3.562	2.634	1.732
Kurtosis	.176	.365	.494	-.175	13.229	7.179	2.681

Note. A score of 0 indicates an attribution of No emotion, and 7 indicates an attribution of strong emotion.

Joy, Relief and Shame have not been attributed at a significant level (≥ 1) to the audio-visual clips presented to participants. Besides, Joy and Relief have a high Kurtosis. These characteristics lead us to exclude these three scales from further analysis.

The highest emotion attribution is for Fear ($m = 1.61$; $SD = 1.69$), followed by Anger, Sadness and Contempt. None of the attributions had a normal distribution.

Means and standard-deviations for scores attributed by participants to each video are stated in Appendix III.

To answer our research questions, the presentation of the results is organised according to the three ways of exploring the emotion-appraisal link, as enumerated in the analysis section: necessity and sufficiency indices, correlations and regression analyses.

Necessity and sufficiency indices

The relation between appraisals and emotions in terms of necessity and sufficiency was calculated through φ indices. An index score of 0 indicates independence of an appraisal attribution from a particular emotion attribution.

Table 6

Φ Indices for Necessity of Appraisals for Attribution of Emotion Labels

	Appraisal Attribution					
	suddenness	goal obstruction	discre-pancy	coping potential	respect internals standards	external standard violation
Emotion						
Anger	0.119	0.238	0.165	-0.122	0.216	0.22
Sadness	0.222	0.241	0.271	-0.177	0.066	0.07
Contempt	0.100	0.263	0.136	-0.091	0.296	0.30
Fear	0.258	0.336	0.264	-0.218	0.101	0.10
Shame	0.124	0.229	0.094	-0.238	0.146	0.15

Note. A score of 0 signifies independence of appraisal and emotion. A score of 1 signifies necessity of an appraisal for the attribution of an emotion.

Table 7

Φ Indices for Sufficiency of Appraisals for Attribution of Emotion Labels

	Appraisal Attribution					
	suddenness	goal obstruction	discre-pancy	coping potential	respect internals standards	external standard violation
Emotion						
Anger	0.086	0.149	0.121	-0.192	0.232	0.23
Sadness	0.145	0.136	0.180	-0.251	0.064	0.06
Contempt	0.050	0.114	0.069	-0.099	0.219	0.22
Fear	0.197	0.222	0.204	-0.362	0.115	0.11
Shame	0.041	0.066	0.032	-0.172	0.072	0.07

Note. A score of 0 signifies independence of appraisal and emotion. A score of 1 signifies necessity of an appraisal for the attribution of an emotion.

No φ index value was equal to zero and none was equal to one. All values were closer to 0 than to 1. The highest values were .336, determining the necessity of obstruction for Fear and -.362, determining sufficiency of coping for Fear.

Correlations

Spearman's rho (ρ) coefficient (bivariate correlations) was calculated for the appraisals scales and emotions scales. Correlations were run for the predicted relation (one-tailed) and also to explore non-predicted ones (two-tailed).

<u>Predicted</u>: All reported correlations are significant at p<.05 (one-tailed).

Table 8

Spearman correlation analysis of Emotion Labels

Appraisal Components	Anger[2]	Sadness	Contempt	Fear
suddenness	0	0	0	+
				.501
goal obstruction	+	+	0	+
	.274	. 654		. 721
relevance/discrepancy	+	0	0	+
	n.s.	.510		. 576
coping potential	+	-	0	-
	-.344	-.201		-.460
violation of external standards	+ -.276	0	+ .452	0
respect of internal standards	- .607	0	0	0

Note. First line indicates the expected sign of correlations, the second the actual Spearman's coefficient for the appraisals scales and emotions scales.

[2] Emotion labels are capitalised to to mark the difference between third party attributions and affective phenomena experienced by individuals and to visually differentiate them from appraisal labels.

<u>Exploratory: All reported correlations are significant at p<.05 (two-tailed).</u>

We now report other, non hypotheses directed, correlations, first between appraisals, appraisals and emotions, and then emotions.

- suddenness is additionally correlated with
  - obstruction (ρ = .718, $p<.001$)
  - relevance (ρ = .900, $p<.001$)
  - Sadness (ρ = .415, $p<.005$)
  - Fear (ρ = .501, $p<.005$)

- goal obstruction is additionally correlated with
  - relevance (ρ = .621, $p<.001$)
  - external standards violation (ρ = .535, $p<.001$)
  - Sadness (ρ = .654, $p<.001$)
  - Fear (ρ = .721, $p<.001$)

- discrepancy is additionally correlated with
  - external standard violation (ρ = .608, $p<.001$)
  - respect of internal standards (ρ = .404, $p<.01$)
  - Fear (ρ = .569, $p<.001$)

- coping is additionally correlated with
  - respect of internal standards (ρ = .658, $p<.001$)

- violation of external standards is additionally correlated with
  - Fear (ρ = .357, $p<.05$)

- Anger is additionally correlated with
  - Sadness (ρ = .363, $p<.05$)
  - Contempt (ρ = .837, $p<.001$)

- Sadness is additionally correlated with
  - Fear (ρ = .800, $p<.001$)

All reported correlations in Figure 2are significant at $p<.05$ (one-tailed).

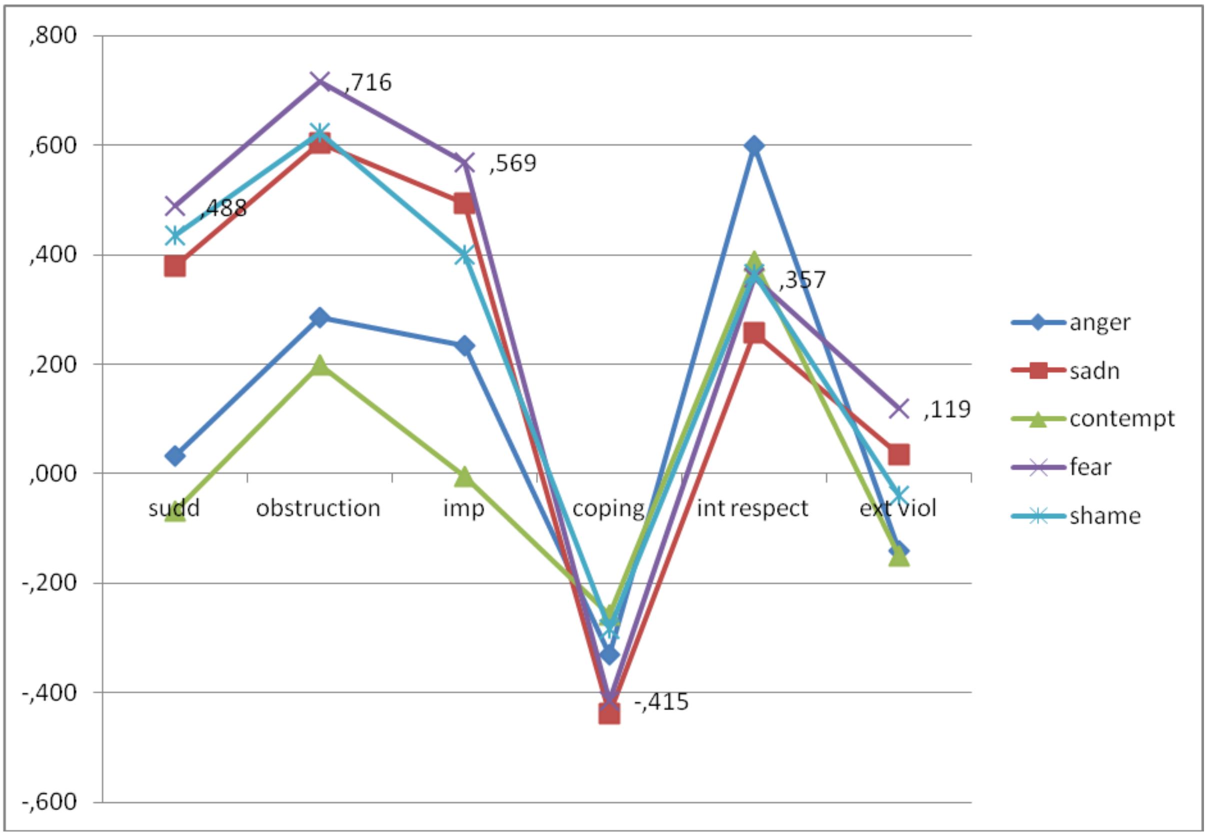


Figure 2 ***Correlation coefficients between emotion scales and appraisal scales***

Regressions

Standard multiple linear regression analysis was carried out to determine the effect of appraisal attributions on emotion attributions. The study was run across all participants. Results are reported in Table 8 on the following page.

Table 9

Multivariate Regression of Appraisals and Emotion Scale Attributions

	Appraisal attribution						
Emotion	suddenness	goal obstruction	discre-pancy	coping potential	respect internals standards	external standard violation	F
Anger	**-.15** **(p<.005**)	**.144** **(p<.005)**	**.121** **(p<.005)**	**-.108** **(p<.005)**	**.365** **(p<.005)**	-.039 (p = 160)	55.20 (p<.005)
Sadness	.013 (p>.05)	.057 (p>.05)	**.244** **(p<.005)**	**-.283** **(p<.005)**	-.021 (p>.05)	.044 (p>.05)	45.09 (p<.005)
Contempt	**-.103** **(p<.05)**	**.162** **(p<.005)**	-.016 (p>.05)	**-.066** **(p<.05)**	**.321** **(p<.005)**	-.027 (p>.05)	33.45 (p<.005)
Fear	.12 (p>.05)	**.137** **(p<.005)**	**.165** **(p<.005)**	**-.317** **(p<.005)**	.01 (p>.05)	**0.84** **(p<.005)**	50.48 (p<.005)

Note. Standardized Coefficients (with p values). Significant results are in bold.

2.10. Discussion

In this chapter we looked at how emotions and appraisals are related, from the perceptive point of view, namely how these two phenomena are judged in other individuals by third parties. To this end, participants were asked to watch short videos of passengers claiming the loss of luggage and to judge the inner state of passengers. Inner state was defined in terms of sentences describing appraisals and emotion labels and evaluated with seven point likert scales.

The nature of the relation between the different components of the inner state was observed in terms of necessity, sufficiency and contingency of appraisals contributing to emotions.
In our results all φ indices are above 0 and below 1. This means that none of the appraisals can be confirmed as being necessary or sufficient for the attribution of any particular emotion. These results confirm our expectations. A person can be perceived as being angry without being perceived as having a goal obstruction, angry without the ongoing event being appraised as important and relevant for the person and without an obligatory attribution of external standards violations, etc. (necessity of particular appraisals is not confirmed). A person can be perceived as having a goal obstruction, or having coping potential without that perception leading directly to the attribution of anger, or the direct attribution of any other emotion (sufficiency of particular appraisals is not confirmed).
As in CPM, there is no statement that all components (cognitive, motor, subjective perception, etc.) are necessarily present for the elicitation of an emotion.

All values being closer to 0 than to 1 tends towards the idea that appraisals and emotions are more independent from each other (*A score of 0 signifying total independence of appraisal and emotion*) than dependent (*A score of 1 signifying necessity of an appraisal for the attribution of an emotion*). However we did not formulate predictions for the range of φ indices to be observed, it is somewhat interesting to note that all results tend towards independence of emotion attribution from appraisal attributions. In our study, the appraisal with the greatest necessity for the attribution of an emotion was obstruction – with a value of.336 in the attribution of Fear. No theory or study result known to us describes obstruction as more important for the triggering, expression or attribution of Fear than of some other emotions, nor is this relation between appraisal-emotion stronger or less ambiguous than in the case for example of Joy and no obstruction.
Surprisingly, the sufficiency index score was also the highest in the case of an appraisal associated with Fear - a negative score for coping potential of -.362. Although the relation between non-coping and Fear is expected, that is that non-coping (the state or the attribution) could be sufficient to lead to fear (the state or attribution) is one to be expected, nothing leads us to believe that this one will be more sufficient than others.

On the contingency level, relations between appraisals and emotions were also studied using correlations and regressions. Our results replicate to a large extent the ones obtained in previous studies.

Results from correlations show that Anger was positively associated with obstruction, as expected.

All other hypotheses regarding Anger were confirmed only partially. The other correlations were also significant, however they did not go in the expected direction. Thus, Anger was negatively correlated with coping potential and negatively correlated with external standard violation. It was also positively correlated with internal standards, and strongly so ($\rho = .607$, $p<.001$), although again expected to be negative according to Scherer (1999). Anger is further positively correlated with relevance and negatively with suddenness.

Our results confirm observations reported by Kuppens and his colleagues (e.g., Kuppens, Van Mechelen, Smits, De Boeck 2003). The results of a series of their studies on anger show that accountability and a sense of “arrogant entitlement” (a combination of respect of internal standards and moral superiority) were specific to anger. The authors accentuate however that no single appraisal was necessary or sufficient.

Kaiser and Wehrle (2001) also report, based on results from their experimental task in which they manipulated appraisal-eliciting situations, that external standard violation is an important appraisal dimension for anger, but it is not a necessary determinant. They differentiate anger, with different appraisal profiles. They find reports of anger in situations that participants judge as being of external standard violation but without blaming any other person. Moreover, participants also report anger in situations appraised as not showing any external standard violation but when an external character was evaluated as “doing it on purpose” (Kaiser & Wehrle, 2001).

Multivariate regressions show that five appraisals could enable to predict Anger attribution: suddenness, obstruction, relevance, coping and respect of internal standards. It appears that respect of internal standards is the best predictor for the attribution of Anger (Beta = .365).

To explain this finding, one can follow Roseman’s argument that how persons evaluate their coping potential may sometimes be based on their perception of “legitimacy”. In our case, we could extrapolate that the situation faced by passengers is in some cases of having „justice on one’s side” – although passengers appraise that they respected internal standards/norms and have acted in accordance with expectations, something negative has happened to them, which is due to external agency. Legitimacy is the perception that although one has done all that was due, one was unfairly treated, in a discriminatory way. Such legitimacy may typically increase perceived control potential, „because legitimacy or deservingness can persuade other people to accede to one’s wishes or to provide assistance in attaining otherwise unreachable outcomes”(Roseman, Spindel & Jose, 1990).

It is necessary to underline that coping potential is not positively linked to Anger, although the correlation is still significant. We could argue, that following different former studies (Frijda et al., 1989; Roseman et al. 1990; Kaiser, Wehrle 2001) we obtain results that do not confirm the CPM trend, in which the perception of control or power over what is happening is considered necessary for the perception and expression of Anger. However this might be due to our formulation of the coping potential statement that uses the terms “being able to master the situation”. Given that one is not able to act directly on finding luggage, the question might not be correctly provided. What should be eventually assessed is the possibility to act at all, that is to have the belief one can complain or state the loss so that luggage can be restored. Such a formulation could in fact stress rather the right to complain.

Thus we might see our assessment of respect of internal standards as answering the need of coping potential stated as contributing to anger in the CPM (e.g. Scherer, 1999), given that it might provide a similar sense of possibility to act. This would explain the high correlation with respect of internal standards.

What is more, Ellsworth and Tong (2006) argue that anger at one-self and at others differ in terms of appraisals. Angry people in their study felt that they were in the right (respect of internal standards), and that what happened was unfair and morally wrong (external standard violation). People who were angry at themselves felt that they were wrong (they did not respect their internal standards), but do not see the situation as involving a moral violation (no external standard violation).

Following that line of thought, it would be interesting to investigate whether participants attributions of anger were reports of passengers' being angry at themselves (as the pattern of appraisal we observe is similar to what Ellsworth and Tong report for self-anger) or angry at other people.

For Sadness, Contempt and Fear, all expected correlations were found. They were significant and in the right direction.

Sadness was correlated with goal obstruction and coping, as expected. Two variables contributed: discrepancy positively (Beta = .244) and coping negatively (Beta = -.283).

Contempt was correlated with external standard violation, as expected. Four appraisals contributed to the attribution of Contempt: suddenness (Beta = -.103), obstruction (Beta = .162), coping (Beta = -.066) and respect of internal standards (Beta = .321). Once again, respect of internal standards is the appraisal that best predicts attribution of the emotion. Thus, what seems to differentiate Anger from Contempt, in terms of appraisals observed, is the attribution of discrepancy which is an appraisal contributing to Anger more than to Comtempt.
Thus, for Contempt we observe a similar pattern as for Anger, except for the appraisal of relevance which is not linearly linked to the former. Contempt is positively associated with obstruction and respect of internal standards, but negatively with suddenness and coping.

To Fear, suddenness, obstruction, relevance, and were positively linked, whereas coping was negatively, as expected. The latter was the best predictor (Beta = -.317).

2.11. Conclusion

We remark that our results are specific to the particular negative situation of passengers losing luggage. Although our data have ecological validity, with several persons observed in a similar situation, results cannot be claimed as perfectly representative of everyday processes, not even of other instances of negative situations.

When looking at means, we see that respect of internal standards has been only slightly attributed. We could hypothesise, that passengers are in a somewhat uneasy situation, in which they have to force upon themselves an unpleasant administrative task. There might be a friction between the value they attribute to their luggage and how petty this loss might seem to the exterior. However we have to keep in mind that participants reported a difficulty in understanding the two questions regarding standards.

Finally, we remark that the filmed passengers acted in a social context, endowed with rules which are commonly accepted and shared in a social community. Many behaviours that occur in social interactions are not simple spontaneous communications of internal states (e.g. see Kaiser & Wehrle, 2001) but are often accompanied by signals communicating also non affective information (Fridlund, 1994; Russell, Fernández-Dols, 1997). Even when reflecting emotions, expressions are not pure read-outs of inner states, as social contexts require to some extent the application of display rules (Ekman, Friesen, 1969). These would for example regulate the expression of negative emotional states by encouraging suppression for general politeness purposes, or in case of Fear suppression of its expression for maintenance of a positive self-image.

We could hypothesise that the communication of anger is hidden to conform to social rules. Wagner and Smith (1991) claim that when the interlocutor is less a co-actor performing the same task than an audience, negative expressions tend to be inhibited.

Keeping in mind all these considerations, we report that our study supports the idea that none of the attributed appraisals is necessary or sufficient for the attribution of considered emotions.

Our analyses confirm the majority of the expectations we based on appraisal theories, in terms of emotion and appraisal correlations. For Anger, our expectations are only partially confirmed. Three of four expected appraisals are correlated, although two of these, coping potential and external standard violation, are negative in our study instead of showing the predicted positive correlation. Anger was positively associated with obstruction, as expected. Respect of internal standards was the best predictor.

We explain these results by making a parallel with Roseman and colleagues’s emphasis of „legitimacy” as a situation leading to anger (1990).

For Sadness, Contempt and Fear, all predictions were confirmed. Sadness has a positive correlation with goal obstruction and a negative correlation with coping potential. Contempt has a positive correlation with external standard violation. Fear has a positive correlation with suddenness, goal obstruction, discrepancy and a negative one with coping potential.

The correlations we find enable us to say that participants do interpret the videos they see in terms of coherent appraisal and emotion patterns. Following this idea, we further analyse the appraisal and emotion label scores provided by participants and look how these correlate with the nonverbal behaviour expressions perceived in videos.

3. Problematics 2: Expressions of appraisal and emotion

L'âme est donc la source de l'expression

Duchenne, 1876

3.1. Introduction: Decoding appraisal and emotion expressions

Nonverbal behaviour, such as face or body movements or postures, can transmit information that is recognised as expressions of internal states or processes, such as cognitive evaluations, cognitive effort, concentration or emotions as described and illustrated in Darwin's "The Expression of the Emotions in Man and Animals" (1872).
Prototypical expressions have been described, with a consensus for six emotions, namely happiness, sadness, surprise, disgust, anger and fear (for review see Ekman, Friesen & Ellsworth, 1982; Scherer & Ellsworth, 2007). They are considered universals, as in the fact that they are believed to be experienced, interpreted, expressed and perceived in others in similar ways independently from cultural backgrounds.
Tomkins (1963) explains the universality of facial expression by their biological basis, while Ekman continues this idea further by claiming affect programs to be "open programs", however somewhat resistant to change (Ekman, 1998). Izard claims that some prototypical expressions are innate neuro-programs that are functional at birth (1994), however this has not been confirmed by developmental studies. Induction studies of surprise in young infants, for example, did not confirm the presence of prototypical facial expression (Camras et al. 2002; Scherer, Zentner & Stern, 2004). Appraisal theorists argue that simply it might be that *the cognitive mechanism for certain appraisals is not yet available* at an early age (Scherer et al., 2004) and does not lead to the associated facial expression.

Thus, although in many research fields related to emotion, as well as in laypeople's perception, the existence of prototypical expressions enabling to differentiate particular emotions is considered self-evident, the mechanisms underpinning these patterns of behavioural changes are unclear (Scherer & Ellgring, 2007).

It is important for affective researchers from different approaches to keep advancing research on emotion-expression interaction, and the emotion-cognition-expression interaction. One appropriate way for such an unpacking seems the use of componential theories of emotion, such as the Component Process Model (CPM) which stresses that an emotional state is the result of various appraisals. Rather than claiming the existence of *a limited number of basic emotions characterized by emotion-specific patterns of expression, this theoretical approach suggests that expressions (and other response components) are direct, efferent products of the results of appraising a stimulus event for specific evaluation dimensions* (Scherer et al., 2004).

While an emotion relies on successive evaluations of a stimulus, it is also composed of different, interconnected and synchronised (motor and physiological) changes that are linked to the sequence. The sequence is defined by a series of checks enabling the evaluation of a stimulus, whether it is internal or external.

In CPM, facial expression changes are associated with appraisal check outcomes. These expressions are described in terms of minimal facial muscle movements, that is facial Action Units (AU) as described by Ekman and colleagues in the Facial Action Coding System (FACS; Ekman, Friesen, & Hager 2002; for a more detailed description see section *Facial Action Coding* 3.5.1, in Problematics 2: Expressions of appraisal and emotion). The direct link between facial expressions and evaluation checks is described through AU, each AU being attributed to a specific outcome of an evaluation check.

According to Scherer and Ellgring (2007) the same emotional label could be attributed to internal states characterised by slight variations in the outcome of appraisal checks.
Facial expression could not always be linked directly to the emotional label, given that one label can be attributed to differentiated states/appraisal outcomes. It is the result of each appraisal check that is hypothesised to be directly related to particular facial movements, while expressions of one emotion would be more diversified.

3.2. Research question

We want to examine if facial expressions are directly linked to mental state attribution by observing the link between seven appraisal attributions, action units (AU) and seven emotional labels. We have chosen the appraisal checks for which concrete facial action predictions have been formulated in CPM theory: 1) suddenness, 2) goal obstruction, 3) relevance (this is novel, important and incongruent with expectations), 4) coping potential (mastery over situation), 5) no coping potential (no mastery over situation) 6) violation of internal standards and 7) violation of external standards (this is unfair and immoral).

We expect to confirm Scherer and Ellgring's predictions (2007) for the five first appraisals. For the internal standards violation, an additional AU is expected (based on Kaiser and Wehrle, 2001) because of a lack of unanimity between researchers who observe different facial changes (Kaiser & Wehrle, 2001).

3.3. Predictions

3.3.1. Theoretical hypothesis

An effect of observed facial and body actions is expected on the perception, and consequently reported attribution, of emotion and appraisal judgment, as apparent in third party ratings of spontaneous behaviour in short video clips.

3.3.2. Operational hypotheses

Hypothesis 1

An appraisal check is expected to be directly linked to the presence or absence of a particular AU or of a set of AU (H1).

In particular, we expect to confirm Scherer and Ellgring’s expectations (2007, see Table 1) concerning the first five appraisals (H1a).

Table 10

Expected Associations Between Action Units and Cognitive Evaluations based on CPM

Characteristics of the event	Facial Action Units expected to be observed
sudden	1 + 2
goal obstructive	17 + 23; 17 + 24
goal Conducive	6 +12+ 25
relevant and discrepant	4,7,23,17, gaze directed
control (high)	4,5 or 17, 23, 25 or 23, 24
(no power)	1, 2, 5, 26, 20, 38

Note. Predictions based on Scherer and Ellgring, 2007, Table 1

Concerning standard violation, we expect to confirm results reported in other studies (Kaiser et al., 2001; Alvarado & Jameson, 2002) with AU 14 (buccinator muscle contraction) for the internal as well as external standard violations (H1b) and no AU 10 for the external standard violation as predicted in the CPM (e.g. Scherer & Ellgring, 2007).

**Action Unit Legend*
1 = inner eyebrow raise, 2 = outer eyebrow raise, 4 = brow lowering, 5 = upper lid raise (wider eye opening), 6 = crows feet/cheek raise, 7 = lower eyelid contraction, 10 = upper lip raise, 11 = nasolabial furrow deepening, 12 = upward lip corner pull (smile), 14 = dimples, 15 = downward lip corner pull, 17 = chin raise, 20 = lip stretch, 22 = lip funneling; 23 = lips tightening, 24 = lips pressing, 25 = lips parting (mouth opening), 26 = jaw drop

Table 11

Expected Associations Between Action Units and Cognitive Evaluations not based on CPM

Characteristics of event evaluated by participants	**Facial Action Units expected to be observed**
respect of internal standards	14
external standards violation	14 and no 10

Note. Predictions based on Kaiser et al., 2001 and Alvarado & Jameson, 2002

Hypothesis 2

An emotion label is expected to be directly linked to the presence or absence of a particular AU. In particular, we expect to confirm Scherer's findings (2001) and Ekman's Emotional Dictionary.

Table 12

Expected Associations Between Action Units and Emotions

Emotions evaluated by participants	**Facial Action Units expected to be observed**
Joy	1, 2, (5), 6 + 12, 26
Anger	1, 2, 4, 5, 6, 7, 10, 23, (24), 25, 26, 27
Relief	no prediction
Sadness	1, 4, 5, 15, 17, 25
Contempt	4, 7, 10, 11, 15, 17, 25, 26
Fear	1, 2, 4, 5, (7), 20, 25
Shame	1, 2, 4, 5, 14, 22, 23, 25

Note. Derived from Scherer (2001) and Ekman's Emotional dictionary (unpublished)
"+" signifies simultaneous AU and "," signifies alternative ones.

Hypothesis 3

For the attribution of six emotion labels (Joy, Anger, Sadness, Contempt, Fear, Shame), we expect to confirm some of the descriptions provided by Darwin (1872) and also observed by Dael and her colleagues in their recent work (2012). While Darwin worked on the link between expression and emotion in humans in observational data, Dael and her colleagues worked on an extensive collection of portrayals of emotions from professional actors guided by standardised emotion-eliciting scenarios (for the corpus description see Bänziger & Scherer, 2010).

Table 13

Expected Associations Between Body Actions and Emotions

Emotions evaluated by participants	Body actions expected to be observed
Joy	Beats, head nods, erect position (no slouching)
Anger	Fist, erect posture (no slouching/upper body is not collapsed), beats and illustrators (communicative gestures), forward lean of torso*,
Sadness	Slouching, diminished number of behaviours (loss of overall muscle tone)
Contempt	Torso turned to the side°
Relief	Decrease in number of behaviours*, slouch*, loss of overall muscle tone*
Fear	Hands alternately clenched and opened with a twitching movement, gaze directed towards and away from the stimulus, backward lean of torso*
Shame	Upper body collapsed°, lowered gaze°, lowered head°

Note. Predictions derived from Darwin° (1872) and Dael (2012).*

3.4. Analyses

We explore the relation between the co-presence of particular face or body units and different attributions. Bivariate correlations were calculated and reported between the mean participants' attributions for each video and the FACS coding scores of each video (N = 41). The annotation in terms of FACS units is quantified through the frequency of AU coded in each video (freq), the total duration of this AU in a video (durT) and the mean duration of this AU (durM), that is the duration of the AU divided by the total duration of the video.
When a simultaneous couple of AU is expected by theory, only the frequency of appearance of both on same frames is recorded and not the common duration.

3.5. Materials

3.5.1. Coding the Lost Luggage Corpus

As we want to associate short extracts to attributions made by third parties, it is important to code all the cues that can have an impact on the observers. For that reason an action or a position is described as happening in a given time, even when we do not see the movement starting from a neutral position, or even when we do not see the movement at all, e.g. when the muscular contraction has already started before the beginning of the video extract and the contraction is kept during the whole clip. Thus, the action units will not be necessarily coded in comparison with a neutral expression of the observed person. Instead, it may happen that a coder codes some stable physical characteristics of the face, if they are marked and exaggerated by nature, in terms of facial AU s or body units.

When we code manipulators, we do not include the preparation and the retraction phases. A unit starts when the action actually starts, e.g. when a hand touches the face to scratch it (and not when the hand goes up to the face); it finishes when the hand stops touching the face. Placing a hand against another surface for no practical reason is considered a self manipulation, e.g. when placing one's head on one's arm, a hand against the cheek.

For the head nods and the shakes, the coding was focused on the duration and not the frequency, that is the beginning and the end of a period during which the actions appeared were coded and not the particular movements.

Pointing to an object was included into the class of illustrators.

Arms crossed implied a position of the arms where each hand was held further than the other. Arms could be lying on the desk, but a simple position of support on the desk, hands clutching each other was not sufficient to score it.

As we want to associate the short extracts to attributions made by individuals, the action units will not be coded in comparison with a neutral expression of the observed person. As a consequence, it can happen that stable physical characteristics of the face, if they are marked and exaggerated by nature, could be coded in term of AU s.
There was no distinction in the code between actions and stable features.

Defining a methodological framework

Coding scheme

The face and body actions were annotated in the Anvil software, version 5.0. beta 12. The script was written in XML. The annotation window provides behaviour code tracks, 61 for the Facial Action Coding System (FACS; Ekman, Friesen & Hager, 2002) and 22 bodily action coding in time (see next section and Appendix IV).

Facial Action Coding

The Facial Action Coding System is a comprehensive system that inventories and classifies visually distinguishable facial movements. It is anatomically based and relies on muscle movements. An action unit is not linked obligatorily to one muscle, but can involve a number of muscles that act together and cannot be dissociated. In other words, a pattern of distinctive features describes each action unit: movements of the skin due to the contraction causing displacement, the characteristic changes in the shape of elements of a face (typically of the eyebrows) and, and the pouching, bulging and wrinkling of the skin.

For FACS coding, the annotation window was divided into: upper face actions, lower face actions, miscellaneous actions, eye actions and head actions.

Head actions were coded entirely through FACS: left, right, up, down, tilted right, tilted left, head nods and head shakes. For head nods and shakes, the coding was focused on the duration and not the frequency, that is the beginning and the end of a period during which the actions appeared were coded and not the particular movements.

Body Action Coding

In addition to the FACS action units, we created additional codes to annotate changes in the body actions and postures. These coded for additional gaze orientation, torso positions, hand, arm and shoulder movements, self-manipulators, illustrators, beats and other body movements (Appendix III). We opted for binary scale for our coding. An action is either present or absent.

Calculating inter-rater agreement

Inter-observer agreement has been assessed with Cohen’s Kappa (k, Cohen, 1960; see Appendix V for the formulas).

The FACS coding was done by a certified FACS coder (coder C) and was verified by a second certified FACS coder (coder F), who annotated 5 videos out of 41 (12 %). The videos to be double-coded were assigned randomly. The FACS manual was used by both coders as a constant reference criterion. A more detailed set of scoring rules has also been established to suit the particularities of the task: temporary degradation of the quality of the video, subtle actions, frozen positions from the beginning to the end of clips, obstructions due to hand movements and glasses. Each video had to be played first at normal speed, before being reviewed in slow motion. When in doubt about one of the AU, the coder was advised to come back to its evaluation a second time after the completion of all other AU from the list and to replay the event up to three times.

Each AU had to be scored one by one, starting from the upper face, to the lower face, finishing with the miscellaneous action units and head and gaze units.

To annotate an action as present, « start » and « end » tags were placed at the onset and offset of the event. At this stage, precise location of the event on the time line was done at a frame-by-frame resolution.

The body coding was annotated by the first coder (coder C) and was verified by a third coder (coder B), who was trained by the first coder and who annotated 8 videos out of 41. A similar set of scoring rules has been established as for the FACS coding and, again, the videos to be double-coded were assigned randomly.

In assessing precision of scoring we looked at the frame-by-frame agreement by computing Cohen's Kappa (k) for face and body action coding (Cohen, 1960). Please see Table 14 for Upper face actions, Table 15 for Lower face action units and Table 16 for Body actions. Following Cicchetti and Sparrow (1981) and Landis and Koch (1977), we consider all coding agreement to be satisfactory except for AU 20, (Table 15) where k in the .21-.40 range indicates a weak/fair agreement.
For the action "slouch" (404) the agreement is of 100 % as it has been attributed for the entire duration of one video clip coded by two coders.
We are not able to report Kappa scores for some of the codes, because of their low frequency of attributions by at least one of the coders. We report in the "*Frames*" column the number of frames where both coders agreed on the presence of an action.

Table 14

Kappa for Upper Face Action Units (frame level)

AU	Frames	Frame level kappa
AU 1	254	.804
AU 2	248	.757
AU 4	111	.443
AU 5	100	.602
AU 7	56	.649

Table 15

Kappa for Lower Face Action Units (frame level)

AU	Frames	Frame level kappa
AU 10	104	.76
AU 14	69	.55
AU 17	202	.847
AU 20	26	.353
AU 23	50	.93
AU 26	149	.513

Note. Actions 15, 16, 23, and 24 were coded b y coder C but not coded by Coder F. Actions 50 and 25 were coded by Coder C and provided to Coder F.

Table 16

Kappa for Body Actions (frame level)

Body actions	Frames	Frame level kappa
400	294	.69
402	523	.45
403	327	.65
404	321	1.00
405	51	.98
406	519	.58
407	659	.65
601	436	.46
602	33	.93
604	14	.90
605	202	.68
606	185	.78

Note. Not enough occurence has been attributed to body action codes 401, 600, 603, 607, 608, 609 and 610 by either of the coders and therefore no Kappa has been computed.

3.6. Procedure

We use the data collected in the *Naturalistic expression study* described formerly (see chapter *Appraisal Components of Emotion*). To read about the procedure used, see section **Erreur ! Source du renvoi introuvable. Erreur ! Source du renvoi introuvable.**.

3.7. Results

3.7.1. Descriptive statistics

We report the statistics of face action units (AU) annotated in our video clips. We provide the number of clips in which an AU appears at least once (clips with AU), the percentage of clips in which that AU appears (% clips with AU), the mean number of times an AU occurs in videos (mean) and the maximum number of times an AU occurs in the videos (max freq).

These statistics for the total duration of each AU and the mean duration of each AU are reported in Appendix VI.

Upper Face actions

AU	1	2	4	5	6	7	43
clips with AU	33	32	25	13	7	29	9
% clips with AU	80.5	78.0	61.0	31.7	17.1	70.7	22.0
Mean	2.3	1.7	1.5	0.6	0.2	1.7	0.3
max freq	7	6	8	8	1	8	3

Lower Face actions

AU	9	10	11	12	13	14	15	16	17	20	23	24
clips with AU	1	17	4	14	3	28	14	6	27	12	23	20
% clips with AU	2.4	41.5	9.8	34.1	7.3	68.3	34.1	14.6	65.9	29.3	56.1	48.8
Mean	0.0	0.9	0.1	0.8	0.1	1.4	0.4	0.2	1.1	0.3	0.9	0.8
max freq	1	4	2	5	1	7	2	2	5	1	7	4

Lips and jaw opening

AU	8	18	22	25	26	27	28
clips with AU	2	2	4	30	26	3	3
% clips with AU	4.8	4.9	9.8	73.2	63.4	7.3	7.3
mean	0.07	0.1	0.1	1.3	1.0	0.1	0.1
max freq	8	3	3	5	5	1	4

Eyes positions

AU	61	62	63	64
clips with AU	15	21	5	34
% clips with AU	36.6	51.2	12.2	82.9
mean	0.5	1.0	0.1	3.1
max freq	3	7	2	10

Head positions

AU	51	52	53	54	55	56	57	58
clips with AU	22	25	17	29	12	9	13	9
% clips with AU	53.7	61.0	41.5	70.7	29.3	22.0	31.7	22.0
mean	1.0	1.0	0.6	1.4	0.3	0.3	0.3	0.2
max freq	7	5	3	7	2	3	2	1

Miscellaneous actions

AU	19	29	30	31	32	33	34	35	36	37	38
clips with AU	0	5	0	1	1	0	2	0	0	1	0
% clips with AU	0	12.2	0	2.44	2.44	0	4.88	0	0	2.44	0
mean	0	0.15	0	0.02	0.02	0	0.07	0	0	0.02	0
max freq	2	0	2	0	1	1	0	2	0	0	1

Gross behaviour scores

AU	50	80	84	85
clips with AU	29	3	3	8
% clips with AU	70.7	7.32	7.3	19.5
mean	1.9	0.07	0.1	0.2
max freq	7	1	2	2

3.7.2. Results

Correlations between attributions (appraisals, emotions) and behaviour coding (frequency, total duration and mean duration of FACS and body units activations) were performed.

Correlations

Given that the participants' results data did not have a normal distribution, a non parametric analysis, Spearman's rho (ρ) coefficient, was computed. Bivariate correlations were calculated on the mean attributions for each video and the FACS coding (AU frequency (freq) and AU total duration (durT) in a video and AU mean duration of this AU (durM) for all the videos)

All reported correlations are significant at $p<.05$ (one-tailed)

suddenness
No correlation with AU 1, 2, nor with the frequency of AU 1+2

obstruction
Positive correlation with AU 4: freq4 ($\rho = .261$, $p<.05$)

Negative correlation with AU 12: freq12 ($\rho = -.269$, $p<.05$), durT12 ($\rho = -.320$, $p<.05$), durM12 ($\rho = -.348$, $p<.05$)

No correlation with the presence of AU 17+23 ($p>.05$) nor with 17+24 ($p>.05$). No correlation with AU 6 ($p>.05$)

relevance/discrepancy: no correlation with AU 4, AU 7, AU 17 nor 23

coping potential
Positive correlation with AU 7: durT7 ($\rho = .373$, $p<.01$); durM7 ($\rho = .378$, $p<.01$). For freq7 no correlation was found ($\rho = .225$, $p = .079$)
Positive correlation with AU 17+23: freq17+23 ($\rho = .261$, $p = .05$). No other predicted correlations significant

respect of internal standards
No correlation was found with AU 10, nor AU 14 ($p<.05$)

violation of external standards
No correlation was found with AU 10, nor AU 14 ($p<.05$)

Anger
Negative correlation with AU 7: durT7 ($\rho = -.322$, $p<.05$), durM7 ($\rho = -.366$, $p<.01$) but not with freq7 ($p>.05$)
Negative correlation with torso forward: freq400 ($\rho = -.368$, $p<.05$), durT400 ($\rho = -.305$, $p<.05$), durM400 ($\rho = -.278$, $p<.05$)
Negative correlation with the slouch: freq404 ($\rho = -.376$, $p<.05$), durT404 ($\rho = -.378$, $p<.05$), durM404 ($\rho = -.383$, $p<.05$)

No correlation found with AU 1, 2, 4, 5, 6, 12, 23, 24, 25, 26, 27 nor with fist actions, erect posture/slouch, beats or illustrators

Sadness
Correlation with AU 1: freq1 ($\rho = .280$, $p<.05$) but not with durT1 or durM1 ($p>.05$)
Correlation with AU 25: freq25 ($\rho = .262$, $p<.05$) but not with durT25 nor durM25 ($p>.05$)
No correlation with AU 5, 15, 17, nor with slouch or diminished frequency of behaviours

Contempt
No correlation with AU 4, 7, 10, 11, 15, 17, 15, 17, 25, 26, nor rotated torso

Fear
Positive correlation with AU 5: freq5 ($\rho = .313$, $p<.05$), durT5 ($\rho = .311$, $p<.05$), durM5 ($\rho = .265$, $p<.05$)

Negative correlation with AU 7: durT7 (ρ = -.281), but not with freq7 nor with durM7 (p>.05)
Positive correlation with self-manipulation at hand-level: freq602 (ρ = .352, p<.05), durT602 (ρ = .441, p<.01), durM602 (ρ = 443, p<.001)

No correlation with AU 1, 2, 4, 20, 25, 26, 27 nor with the frequency of AU 1+2, or with nervous fidgeting, slouch, rotated torso, backward lean of torso

3.8. Discussion

In this chapter we looked at observers inferring emotions and appraisals from nonverbal cues, namely face and body actions they see in audio-video clips. How do laypeople decode expressions? How is this decoded information organised? We dealt with these questions by performing a correlation and regression analysis on appraisals, emotions and expressions.

Our results confirm some predictions already observed in the literature concerning the relation between facial actions and appraisals. Moreover, we found that there is a sizeable correlation between the frequency and duration of facial expressions and emotion attribution (with the only exception of "Contempt" label).

We summarize in the following the relations between expressions and appraisals that emerged from our *Naturalistic expression study*.

For the coping potential appraisal a positive correlation was found with AU 7 (lower eyelid contraction) and with AU 17+23 (raised chin and pressed lips), as expected. Moreover, a positive correlation was found with the head nodding movements.

For goal obstruction a negative relation was found with the smile, AU 12, as expected, but no other predictions were confirmed.

None of our hypotheses concerning suddenness, relevance/discrepancy and the two norm standard appraisals were confirmed.

For Anger, a negative correlation was found with AU 7 (lower eyelid contraction). No other prediction was confirmed for this emotion.

For Fear, a positive correlation was found with AU 5 (opening of the eye) and a negative one with AU 7 (lower eyelid contraction).
For Sadness, there was a positive correlation with AU 1 (inner brow raise) and AU 25 (opening of the mouth).

On the level of affect perception from body cues, some expectations were confirmed in terms of emotion-specific behaviours.

Our analysis demonstrates that some facial and bodily cues are more related than others to mental state attributions by third parties, whether emotion or appraisal attributions.

Specifically, results guide us to wonder if some atomic expressions have an influence on the overall state attribution. One could claim that the mere increase in the activity (whether frequency or duration) of some face or body parts is associated to a change (increase or decrease) of some attribution. This is surprising with respect to non componential emotion attribution approaches, like the basic emotion approach, which suggest that the appraisal/emotion attribution depends on the *whole* expression, which accentuates the importance of a holistic image. Indeed, several emotion theories – which assert specific neuroprograms responsible for triggering expressions as a consequence of emotions – do not claim any clear association with one specific cue. Even in the basic emotion stance, which introduced the idea of "open programs", one does not expect such clear associations.

We remark that correlations do not show more than contingencies of mental attributions and expressions, as present in the considered corpus. The use of correlations gives rise to attributions biased by all kind of interrelations between face and body actions, as some would be naturally occurring more often with some others- we do not explore the impact of the activity of one behaviour while keeping other factors stable To overcome this limitation, in the next chapter we consider an experimental study, in which we manipulate one cue at a time through the use of a virtual character.

3.9. Conclusion

To conclude this chapter, we consider the *Naturalistic Expression study* to have successfully shown correlations between what participants attribute in terms of appraisals and emotions and behaviours observed in videos. These results enable us to build a second study, in which behaviours would not be simply contingent but manipulated one at a time – the *Manipulated Expression study*, that we describe in the following parts of the thesis.

The *Naturalistic Expression study* raises two open issues which need to be further investigated in future research. One is the relation between stereotypical representation of emotions and actual mental attributions by third parties. How could we explain that such stereotypical associations like suddenness, linked to eye brow raise, have not been stereotypically seen enough so as to lead to a positive correlation? Eye brow raise is universally considered a typical expression for suddenness (e.g. Darwin, 1872/1998). Could one reason that the link is not to be spotted through correlations and regressions? Although recently appraisal researchers (Tong et al., 2009; Kappas, 2001) started exploring non linear relations between emotions and appraisals and nothing leads us to think that the link with expressions has to be strictly linear, such an explanation does not seem to be logical in that particular case at least. Another argument is that some facial movements are not displays of emotions, but have communicative function in interaction. Especially eyebrow movements have been observed to have strong illustrative functions in dyadic interactions (Beneke, Merten, Krause, 1998) rather than emotional displays.

The second question we pose is: given the well proven one-to-one mapping between some expressions and some mental attribution, how do we explain the rest of the cases?

Predicted relationships that failed to reach significance could be potentially a result of the need for interrelated behaviours, which cannot be shown with simple bivariate correlations. One behaviour may not lead on its own to a particular state (appraisal or emotion), but may require the presence, absence or certain level of intensity of another behaviour.. Thus one could claim that the cue as such, independently of any other behaviour, or any characteristic of this cue, does not have enough of an impact on the attribution so as to be observed based on participants attribution score. Following Tong and his colleagues (2009) and Roseman and his colleagues (1984), one could claim that emotions are in general to be predicted by combinations of appraisals rather than by individual appraisals. Other parameters have also to be kept in mind when observing the attribution of emotional labels to third parties. It might be that the perception of some appraisals would only lead to an emotion label attribution when the behaviours or the appraisals are in an appropriate sequence and timing and, as basic emotion theories expect it – when contributing to a holistic image.

A more probable reason is that the predicted expressions are not exclusively linked to one mental state, but appear in several ones. It could be eventually possible to measure it when taking in count the intensity of actions, as could be done with the FACS intensity evaluation (scale from A – trace action, to E – maximum intensity of action). For example raised eyebrows are expected in the majority of emotional states, although in much less high intensity than in Fear and suddenness.

4. Problematics 3: Manipulated expressions of appraisal and emotion

...it could be argued that the face has the only skeletal muscles of the body that are used, not to move ourselves, but to move others.

Smith & Scott, 1997.

4.1. Introduction: Fine emotional expression tuning with a virtual character

In former chapters we rely on a video corpus of interacting humans for our perceptive study of internal states (appraisal, emotion) and behaviour (face, body). In our *Naturalistic expression study*, we tackle our first and second problematics through correlational studies. While the results regarding the first problematics confirms us in the idea of contingency between appraisal and emotion attributions by third parties, the results of the second shows contingencies between some behaviours and particular attributions. To establish that the observed relationship is causal, with nonverbal behaviour having a direct effect on attributions of appraisals and emotions, an experimental study was designed. This *Manipulated expression study* enables the isolation of behaviours, for the identification of highly specific movements that have an impact on internal state attribution to third parties.

A novel approach to test the impact of particular behavioural cues on perception by laypeople was to reproduce the human behaviour by a virtual character with many of its subtleties and with its complexity in terms of sequence and use of facial and body movements.

Other systematic methods have been previously used to investigate the contribution of body action units to the expressions of emotion through the use of generated well controlled virtual characters. Thus, Coulson (2004) generated different postures by systematically modifying the degree of rotations of different body segments. The generated postures were shown to participants to investigate how people attribute emotional states to these postures. Roether et al. (2009), on the other hand, focused on the perception of gait. He carried out a three-step process to extract, validate and confirm the minimum set of spatio-temporal motor primitives that drive the perception of particular emotions when watching walking behaviour.

In our study of nonverbal behaviour, we used the virtual character Greta (Bevacqua et al., 2010), developed at Télécom ParisTech, to display tbehaviours seen in the naturalistic corpus described in former chapters. The contribution of specific facial action units was checked by removing them from the reproduced patterns of nonverbal behaviours.

Greta has the possibility to generate face expressions based on FACS, as well as torso and other body movements. However, to enable the display of face and body movements as in the chosen emotional extracts from the corpus, the capabilities of Greta character (Bevacqua et al., 2010) had to be developed to include presentations in sequences of units, with different starting and ending times for particular face and body cues, and superposition of different units.

Thus, work was realised to create a novel approach to the generation of emotional displays in a virtual character. Expressions displayed are not limited to the face, but can be displayed using different modalities, including torso and arm movements. Besides, cues contributing to the expression can occur in sequences. This approach allows for high flexibility and variation of displays which is necessary for the use of Greta character in our perceptive study.

Once the behaviour generation capacities were modified in Greta, leading to the creation of the Multimodal Sequential Expression (MSE) model, evaluation studies were run to verify three main features of our approach, which are the multimodality, sequentiality, and the use of constraints. The results of our first study show that the recognition of the MSE animations is high. The second study enabled us to further observe that multimodal sequential expressions are better recognised than static emotional displays in their apex and (at least for some emotional states) better than dynamical single signals. It also showed that the application of constraints increased the believability of the multimodal sequential expressions.

4.2. Generating behaviours with a virtual character

Studies have shown that human interactions with virtual characters are similar to those developed with real humans (Schilbach et al., 2006; Brave et al., 2005). This enables to hypothesise not only that a human emotion model can be used to model artificial emotions for characters, but also that virtual humanoids can be used as tools in the study of human perception.

Today, several research teams work on the elaboration of embodied virtual characters, also called virtual agents, with interaction capacities, some of which have humanoid appearance. Some of the latter can express emotions or other internal states through their nonverbal cues, mostly facial.

At the beginning of this thesis Greta allowed the generation of expressions from onsets to offsets, with behaviours starting and ending at the same time. Given the believed importance to have sequentiality and superposition of facial and/or body cues (Scherer & Ellgring, 2007; With & Kaiser, 2011), a new approach to the generation of emotional expressions was proposed, developed and to some extent tested.
In comparison to other existing solutions to behaviour movement generation with virtual characters, our Greta character system is equipped so as to generate a variety of multimodal emotional expressions automatically. We build on observational data and in our approach the observed behaviours are described by a human, i.e., by a Facial Action Coding System (FACS) certified expert. The sequences of nonverbal displays are independent behaviours. The system allows for the synthesis of any number of emotional states and is not restricted by the number of modalities.

The approach allows a virtual character to display multimodal sequential expressions (MSE), i.e. expressions that are composed of different nonverbal behaviours partially ordered in time and belonging to different nonverbal communicative channels. Few models have been proposed so far for creating dynamical multimodal expressions in virtual characters (e.g. Pan, Gillies, Sezgin, Loscos, 2007; Lance, Marsella, 2007). More often characters use only stereotypical facial displays which are defined at their apex and then interpolated.

Instead our model generates a variety of multimodal emotional displays of an arbitrary duration. Each of them is composed of a sequence of nonverbal behaviours that are displayed not only by face but also with the use of other modalities like gaze, gesture, head and torso movements. With MSE the repetitiveness of the emotional expressions is avoided by introducing diversity in the signals choice, order and timing. This variability is obtained by probability of appearance and temporal constraints which are defined separately for each signal. In our model a high-level symbolic representation of the behaviour emotional displays are generated from samples described in literature and from annotated videos. Thus, captured data is not directly reproduced, but different plausible expressions of emotions are generated. They are composed from the same signals as the original ones.

Our approach is based on the idea that emotions are processes developing in time. The sequentiality of behaviours in emotional expressions is postulated in the Component Process Model (Scherer & Ellgring, 2007). So far predictions and empirical studies of this componential approach have mostly exclusively described sequences on facial level, however changes in the whole organism are postulated, including other expressive modalities of expression. Other psychologists (e.g. Keltner, 1995; Shiota, Campos, Keltner, 2003) also showed that in the case of some emotions, like embarrassment, sequences can be observed in nonverbal expressions. Again, the face is in action, but other modalities are also contributing to the general interpretation of emotional expressions, interacting with the message transmitted by the face (see chapter *Decoding appraisal and emotion expressions*). Thus emotional expressions may be composed of several multimodal behaviours that do not have to occur simultaneously, but in a sequence that is not due to chance. Keltner described that in the embarrassment sequence (Keltner, 1995) some temporal relations between signals were observed that may be represented in the form of constraints. Other, more recent research on sequences in expression has been realised on the level of the face by With and Kaiser (2011). These authors assert that the coordination in time of facial action units, as well as head and gaze movements, could have an impact on the meaning attributed to these particular cues. They use T-patterns detection algorithm (Magnusson, 2000, 2006) and pinpoint some recurrent temporal patterns of behavioural cues in the studied emotions of enjoyment, embarrassment, hostility, surprise and sadness. The value of using the T-pattern software is that this algorithm enables to see hidden patterns, i.e. those that could be missed by the eye and not to be observed with traditional linear analysis. Some of these patterns are specific to particular emotions.

These findings confirm us in our approach to emotional expressions, and we emphasise the temporal sequences we introduce into affective displays, that go beyond the existing models that focus mostly on facial expression. Introducing the sequences of signals, we aim at enlarging the set of subtle emotional states that can be communicated by virtual characters.

4.2.1. Facial expression models of emotion in virtual characters

Several models of facial expressions have been proposed to enrich virtual character's facial behaviour.

Two different approaches are usually used to create emotional expressions in virtual characters: the motion capture-based and the procedural one. The first one is often used in commercial applications, e.g., in the movie industry. In this approach, the synthesised expressions are characterised by a very high level of details and a great realism. This approach is, however, very time and resource consuming. It may also lack some flexibility and variability - two important issues in an character's behaviour synthesis. In the second approach, an emotional display is generated from a symbolic description. This description is used to define the keyframes of the animations. An interpolation is used to generate a complete animation from particular keyframes.

Usually, a facial expression is presented in its apex (maximal intensity moment is defined as a keyframe), while the animation is interpolated for the rest of frames.

In this approach animations can be of any arbitrary duration, interpolations being run for any number of keyframes spread out in time. However, the most often used models have a trapezoid shape onset-offset-apex. Using such an approach, with three keyframes, generally leads to the generation of a schematic and stereotypic animation.

The existing solutions usually compute “basic emotions” or new expressions by “averaging” the values of the parameters of the expressions of the “basic” emotions (Ekman and Friesen, 1975; Ekman, 2003b).
The model called Emotion Disc (Ruttkay et al., 2003) uses a bi-linear interpolation between two basic expressions and the neutral one. In the Emotion Disc six expressions are spread evenly around the disc, while the neutral expression is represented by the centre of the circle. The distance from the centre of the circle represents the intensity of expression. The spatial relations are used to establish the expression corresponding to any point of the Emotion Disc.

Two models (Tsapatsoulis et al., 2002; Albrecht et al., 2005) rely on the use of expressions of two “neighbouring” emotions to compute the facial expressions for non-basic emotions. For this purpose they use different multidimensional spaces, in which emotional labels are placed. In both approaches new expressions are constructed starting from the six Ekman's expressions: anger, disgust, fear, happiness, sadness, and surprise. To be more precise, in Tsapatsoulis and colleagues (2002) work a new expression is generated by looking for the spatially closest two basic emotions as defined within the dimensional space proposed by Whissell (1989) and Plutchik (1980) and weighting the parameters of these expressions with their coordinates. Albrecht et al. (2005) proposed an extended approach. The authors use a three dimensional space of emotional states defined by activation, evaluation, and power as proposed by Cowie and his colleagues (1999) and anatomical model of the face is used. As a consequence, they work with a numerical representation of muscle contradictions.

Bui (2004) uses a set of fuzzy rules to determine the blending expressions of six basic emotions based on Ekman's findings (Ekman, Friesen, 1975). A subset of rules is attributed to each pair of emotions. The fuzzy inference determines the degree of muscle contractions of the final expression as a function of the input emotion intensities. Arya and his colleagues (2009) propose another perceptively valid model for expression blending. From perceptive study results they develop a set of fuzzy rules that link specific facial actions with the 3D space of valence, arousal and agency. Rules are generated from the statistical analysis of the images created in the experiment by participants who were asked to illustrate short stories with blending expressions. Contrary to Bui whose fuzzy rules were activated depending on the intensity of emotions, in Arya and his colleagues work the fuzzy values in three emotional dimensions are used to activate the avatar face.

Roesch, Tamarit and Reveret (2010) on the other hand, focus on generating facial expressions that can be easily controlled in terms of FACS Action Units (AU; Ekman, Friesen & Hager, 2002). The tool they created, FACSgen, allows users to define in time the dynamics of activation of each action unit separately.

Sequenced expression models of emotions in virtual characters

Pan and his colleagues (2007) proposed an approach to display emotions that cannot be expressed by static facial expressions but that are expressed by certain sequences of signals (facial expressions and head movements). First of all, certain sequences of signals were extracted from a video-corpus. From this real data, Pan et al. built a directed graph (called a motion graph) in which the arcs are the observed sequences of signals and the nodes are possible transitions between them. Different paths in the graph correspond to different expressions of emotions. Thus, new animations can be generated by reordering the observed displays.

Paleari and Lisetti (2006) and Malatesta and colleagues (2007) use manually defined sequential expressions inspired by Scherer's appraisal theory (Scherer, 2001). They consider a limited number of emotions and put the emphasis on the temporal relations between the different dynamic elements of an expression and their link to the consecutive stages of cognitive evaluations. Facial expression is not activated at once; that is, it is not a full blown expression, rather the expression evolves through time. The animation parameters are activated sequentially. The final result is an animation of a sequence of several micro-expressions of cognitive evaluations.

In the work realised by Malatesta and his colleagues (2007), the expressions of anger, disgust, fear, joy and sadness were generated manually according to Scherer's predictions and the focus was on the intensities and on the temporal constraints of facial signals. This work differs from Paleari and Lisetti's work (Paleari & Lisetti, 2006) where each expression is derived from the addition of a new AU to the former ones.

Ruttkay (2001) proposed a system that allows the human designer to modify a facial expression animation defined per default by a trapezoid attack-hold-delay. The system permits, for any single facial parameter, to define manually the course of the animation. The plausibility of the final animation is assured by a set of constraints. The constraints are defined on the key-points of the animation and concern facial animation parameters. One can, for example, force the facial expressions to be symmetric (i.e. all facial parameters have identical values for each key-point). In Stoiber and his colleagues' work (2009) another interface for the generation of facial expression of an avatar is proposed. Using its 2D custom control space the user might deform both the geometry and the texture of a facial model. The approach is based on principal component analysis of the images database showing the variety of facial expressions of one subject. It allows generating both realistic still images as fluent sequence of expressions but deprived psychological ground.

Clavel and her colleagues (Clavel, Plessier, Martin, Ach, Morel, 2009) found that the integration of the facial and postural changes affects users' perception of basic emotions. In particular, an improvement of the emotion recognition was observed when facial and postural changes are congruent (Clavel, Plessier, Martin, Ach, Morel, 2009). Nevertheless, only some models for multimodal emotional expressions have been created so far.

Lance and Marsella (2007) model head and body movements in emotional displays using the PAD-dimensional model. A set of parameters describing how the multimodal emotional displays differ from the neutral ones was extracted from the recordings of acted emotional displays. Consequently, emotionally neutral displays of head and body movements are transformed to multimodal displays expressing, e.g., low/high dominance and arousal.

It is in continuation of such componential approaches that the present advancement of character Greta is situated. In the framework of this PhD project, expressive capacities of Greta had to be improved to enable the display of behaviours observed in the naturatlistic audio-visual corpus, described in *Problematics 1: Appraisal components of* emotion. Body and face behaviours had to be reproduced in sequences by the character and to be coordinated in terms of different modalities, so we have created our model of multimodal sequential expressions (MSE).

In the following sections of this chapter, we present the results of an evaluation of three main features of MSE: multimodality, sequentiality, and constraints. First of all, we check if the character that uses MSE is able to communicate its emotional states properly, i.e., if its multimodal sequential expressions are recognised by humans. We also examine whether using multimodality and sequentiality influences the recognition rate. Finally, we verify the importance of constraints in the perception of believability of the character's behaviour.

In the following sections of this chapter, we present the MSE model. This MSE development description is structured as follows: The next section is dedicated to an overview of computational models of face expressions, with a focus on those that include multimodal and/or sequential expressions. Next, our approach is explained, and then results of evaluations studies of MSE are presented.

4.2.2. Greta character and Multimodal Sequential Expressions (MSE)

The main task of our model is to define displays that could be generated by a virtual character to express particular emotional states. Thus, we created a module in our virtual character system that enables to generate behaviours appropriate for specific emotions. The selection of an emotion label from a set of predefined sates activates what we called the MSE algorithm. The particularity of that algorithm is to generate multimodal sequential expressions of emotions, i.e., expressions that are composed of different signals partially ordered in time and which involve different nonverbal communicative channels.
Our model is based on the following criteria:

- emotional displays are sequences of behaviours on different modalities (see Dael et al., 2011);
- there is variability in the created animations, as one emotion can be expressed through more than one set of predefined behaviours (see Scherer & Ellgring, 2007);
- sequences may have an arbitrary duration;
- animations and sequences are not predefined but are created dynamically, to answer the need of virtual characters to be able to react instantaneously to the changing context and adapt to users in the application of such characters in conversational characters (see Bevacqua, Sevin, Hyniewska & Pelachaud, 2012; Cassell, Geraghty, Gonzalez, & Borland, 2009).

In the following sections, we present details of our approach, starting from observation to the synthesis of emotional expressions with the virtual character.

Data Collection

We base our work on observational studies of human emotion (Haidt, Keltner, 1999; Keltner, 1995; Shiota, Campos & Keltner, 2003; Rozin & Cohen, 2003), as well as on the annotation realised in our laboratory on nonverbal behaviour coming from high intensity emotion displays.

Videos from the EmoTV corpus (Abrilian, Devillers, Buisine, Martin, 2005), the Belfast Naturalistic Emotional Database (Douglas-Cowie, Campbell, Roach, 2003), and the HUMAINE database (Douglas-Cowie et al. 2011), as well as some extracts from live French TV shows, have been chosen in order to observe behaviour expressed in highly emotional situations by non-actors. An annotation scheme was developed to describe face and body cues, as well as emotional states. The extracts have been annotated by a certified FACS coder (Ekman, Friesen & Hager, 2002), to describe visible facial muscular activity, with two to six video extracts per state. For annotating other nonverbal behaviours such as hand, arm, and torso movements, a free textual description was used. An emotional label was attributed in each extract, based on observed expression and the context, e.g., a woman describing the happiest day of her life and using vigorous movements was labelled as cheerful. Although only a very short extract (between 4 and 50 seconds) was annotated, limited strictly to the emotional expression, a longer part of the video clip was viewed to enable comprehension of the context. A detailed description of our annotation can be found in (Niewiadomski, Hyniewska & Pelachaud, 2009).

The behaviour and constraint sets for pride, embarrassment, and anxiety were defined from the literature. The sets of the other five emotional states, anger, cheerfulness, panic fear, relief, and tension, were based on the annotation study (Niewiadomski, Hyniewska & Pelachaud, 2009).

MSE Language

To go beyond characters showing simply static facial expressions of emotion (i.e., expressions at their apex), at first we gathered information on the face and body cues involved in the emotional expressions as well as on the temporal constraints regulating them. Consequently, we have designed a scheme that is based on these observational studies.

For the purpose of generating multimodal sequential expressions, we define a new XML-based language in two steps: a behaviour set and a constraint set. Single cues like a smile, shake, or bow belong to one or more behaviour sets. Each emotional state has its own behaviour set which contains cues that might be used by the character to display that emotion. The relations that occur between the cues of one behaviour set are more precisely described in the constraint sets. The appearance of each cue in the animation is defined by a starting time and an end time.

The behaviour set contains a set of cues of different modalities, e.g., head nod, shaking-hand gesture, or smile, to be displayed by a virtual character. Let us present an example of such a behaviour set. Keltner (1995) described a sequence of signals in the expression of embarrassment. The behaviour set based on Keltner's description (Keltner, 1995) of embarrassment may contain signals like head movements (head up, head down), three facial expressions: smile, tensed smile, and neutral expression, open flat hand on mouth gesture, and a bow torso movement.
A number of regularities occur in expressions that concern signals duration and their order of displaying. Consequently, for each signal in the behaviour set, one may define the following five characteristics:

- Probability of occurrence at the beginning of a multimodal expression;
- Probability of occurrence at the end of a multimodal expression;
- Minimum duration of the signal (in seconds);
- Maximum duration of the signal (in seconds);
- Possibility that the signal might be repeated.

For instance, in the embarrassment example, the signals head down and gaze down occur much more often at the beginning of the multimodal expression than later (Keltner, 1995).

Appearance constraints describe more general relations between signals like inclusion or exclusion, e.g., “signals s_i and s_j cannot co-occur”.

4.2.3. Evaluation of MSE emotional displays

We carried out two studies to validate the MSE approach (for details on the algorithm see Niewiadomski, Hyniewska & Pelachaud, 2011) to the generation of emotional displays for a virtual character. In the first study, we checked whether individuals are able to recognise the emotions expressed by the character. Then, in the second study, we verified if the multimodal sequential expressions are more recognised than static images of emotional displays and dynamical single signal emotional expressions. In the same evaluation, the role of constraints in the perception of multimodal sequential expressions was also checked.
For the purpose of these studies, eight emotional states were chosen: anger, anxiety, cheerfulness, embarrassment, panic fear, pride, relief, and tension. This choice is motivated by the following:

Condition 1. We want to differentiate between several positive emotional states. Usually in the literature, all the positive emotions are described with the general label “joy” and are associated with the raise of lip corners (AU 12) along a raise of the cheeks and crow’s feet around the eyes (AU 6) (Ekman, 2003). In this study, we evaluate: cheerfulness, pride, and relief.

Condition 2. We want to differentiate expressions in which different types of smiles might occur (with and without AU 6, that is cheek raise and crow’s feet). Smiles are used to display positive emotions (e.g., in joy), but they also occur in negative expressions like embarrassment or anxiety.

Condtion 3. We also want to differentiate negative states to be used by the virtual character like anxiety, tension, panic fear and we want to compare them with the expression of anger.

Both MSE evaluation studies have a similar setup. Participants accessed the evaluation studies through a web browser. Each study session was made of a set of web pages, each page presenting one question. The participants could not come back to the preceding question and they could not jump to the question without providing an answer to the current one. No time constraint was put on the task. The questions were displayed in a random order, the emotional labels were ordered alphabetically. Videos were attributed to participants in random order. Participation in studies was anonymous.

MSE study 1. Recognition of MSE Emotional States

First, we were interested in examining if the emotional states expressed with multimodal sequential expressions are recognised by the participants. For this purpose, we showed the participants a set of animations of the Greta character displaying the eight emotional states and we asked them to attribute to each animation one emotional label.
In this study, our hypotheses were the following:

H1.1 Each of the intended emotions is more often correctly recognised on the corresponding animation than chance level.
H1.2 For each animation, the proper label is attributed more often than any other label.
H1.3 We were also interested in the habituation effect, i.e. if showing the same set of animations more than once influences the recognition rate.

Procedure of MSE study 1

Eight animations presenting different emotional displays were used in the study. Participants were asked to recognise the emotions displayed by the virtual character. Each video shows the character displaying one emotional state. The character is not speaking. The duration of each video is about 10 seconds.

After watching an animation, the participants have to attribute one emotional label to the perceived emotional state from an 8-element list before they can pass to another page with a new animation. Participants were told that they could use each label more than once or not at all.

Each study session consists of seeing the same set of eight videos twice presented in a random order. Each subject has to see all eight videos (turn 1) before seeing any of them for the second time (turn 2). They cannot replay the animation.

Results of MSE study 1

Fifty-three participants (25 women, 28 men) with a mean age of 28 years, mainly from France (21 %), Poland (21 %), and Italy (15 %), took part in the study.

None of them works in the domain of virtual characters.

The attribution of correct answers (number of hits) for each emotional expression in both turns is above chance level (which is 12.5 %). In each turn, the greatest amount of hits was for the emotion of Anger (93 % both turns mean), while the least correctly attributed was Embarrassment (41 % both turns mean). The number of hits versus alternative answers in turns 1 and 2 was compared using analysis of variance (ANOVA) and the improvement was not significative (univariate ANOVA, $p>.05$). Therefore, although the analyses for each of the two turns have been realised, the means for both turns are stated for reference in the text when not otherwise

In general, the proper label was attributed more often than any other label. For the animations of Anger, Cheerfulness, Panic Fear, and Relief, the correct labels were significantly more often attributed than any other ones in both turns (McNemar test, $p<.05$ in each turn). For the remaining animations of Anxiety, Embarrassment, Pride, and Tension, the proper label was found but some confusions occurred.

The strongest confusion occurred between Anxiety and Embarrassment. For the Anxiety animation, the number of attributions of the Anxiety (43 % both turns means) and of the Embarrassment (36 % both turns means) labels did not differ significantly (McNemar test, $p>.05$). In the Embarrassment animation, Embarrassment (41 % both turns means) was confused with Anxiety (36 % both turns means) ($p>.05$). In turn 2, Embarrassment (40 %) was also labelled Tension (28 %) ($p>.05$) (while, in turn 1, it was labelled tension by 17 %). Although on the limit of a significant difference ($p = .066$), some other confusions were found: Pride (45 % both turns means) was labelled Relief (26 % both turns means) in both turns and Tension (49 %) was labelled Embarrassment (25 %) in turn 2.

As it might be argued (Russell, 1994) that a correct recognition of a particular emotional expression may not be considered only in terms of correct attributions of a label, but also of rejections of that label for expressions not related to that emotion, we also calculated the unbiased hit rate. For this purpose, we use a

$$k = \frac{(h+cr)-r_{exp}}{(i*j)-h_{exp}} \quad (1)$$

Kappa (k) score, as outlined by Isaacowitz and his colleagues (Isaacowitz et al., 2007): where h is number of hits, cr is number of correct rejections, r_{exp} - chance expected number of responses, i - presented items, j - number of judges, and h_{exp} - chance expected number of hits. K value may vary from 0 (in the case of totally aleatory attribution) to 1 (if a label was always correctly attributed and correctly rejected, i.e., absence of false alarms).

Kappa value was calculated for each emotion. It was satisfactory for all emotions when the eight labels were counted in, with the highest for Anger (0.870) and the lowest for Embarrassment (0.702). Indeed, Embarrassment also had the lowest hit rate (41 %) and the greatest number of false alarms (17 %), showing a general tendency to attribute this label more often to our character's behaviour than any other label. The incorrect attributions of embarrassment were aimed at negative emotions other than Anger (Anxiety, Panic Fear, and Tension).
Since "false alarms" are more likely to occur between similar emotions, we also compare each emotion (summed attribution form the two turns) against the others from each condition, C1, C2, and C3. In C1, each of the three emotions was compared against two more labels. Relief had the highest unbiased score ($k = 0.503$), then Cheerfulness ($k = 0.494$) and Pride ($k = 0.356$). In C2, each emotion was compared against four others: Cheerfulness and Relief had the highest recognition ($k = 0.697$), then Pride ($k = 0.671$), Anxiety ($k = 0.548$), and Embarrassment ($k = 0.513$). In C3, against four other labels, Anger was most recognised ($k = 0.807$), then Panic Fear ($k = 0.715$), Tension ($k = 0.612$), Anxiety ($k = 0.558$), and Embarrassment ($k = 0.547$).

Discussion of MSE study 1

The main aim of this evaluation study was to check if the multimodal sequential expressions are recognised by the participants. The hypothesis H1.1 was verified: The simple recognition rate (41-93 % both turns means) exceeds strongly chance level and the unbiased hit rate measured by is also satisfactory when the chance level is brought to in-group comparison instead of all the eight labels. The hypothesis H1.2 was only partially verified: Although the number of attributions of correct labels was higher than that of alternatives, the difference was not significant for some emotions. Finally, we observed that the effect of habituation (hypothesis H1.3) is not significant and, consequently, multimodal sequential expressions may be used straight-away in short period interactions with the user.

While the recognition rate is quite high, we believe it could have been higher if behaviour expressivity was considered. In the videos used for this perception study, emotions were conveyed through signals defined in the behaviour set. Behaviour execution did not vary, that is, behaviours had the same expressive qualities in all the videos. However, body expressivity is an important cue to convey emotional states, as Wallbott (1998) claims and as we can infer from our corpus annotation. The non-adaptation of the behaviour expressivity to the particular states might have influenced their perception and might have created a general bias. For instance, it appears that participants have a higher tendency to attribute embarrassment when judging the behaviour of our character, particularly when the emotional expressions have a negative aspect and do not portray anger. Thus, we believe that our MSE model should be extended in the future by a number of expressivity features. The emotions that received the highest recognition rate - anger, cheerfulness, panic fear, and relief - are those that are described by facial expressions as well as by specific body and arm movements (e.g., anger with the hands on the hips and cheerfulness with raised arms). It might be that expressions of emotions that make use of the full body were better perceived compared to expressions of emotions conveyed mainly with the face (such as embarrassment and tension). However, this effect may also be explained by the framing used in this study. The animations showed half body of the character and, consequently, the face was quite small. The use of multimodality in communicating emotions should be more carefully analysed (for example, by studying how each modality contributes to the recognition of the internal state).

Nevertheless, our results show that even such subtly differentiated expressions like those of relief or of cheerfulness were recognised surprisingly well. One could argue that none of these expressions probably could have been recognised from still facial expressions in their apex or dynamical single signals, such as a hand or gesture movement. This claim is checked in the second evaluation.

MSE study 2. Role of Sequentiality, Constraints, and Dynamical Signals

In the previous section, we showed that the emotional expressions generated with our algorithm are recognised.

We also suggested that the MSE might be particularly useful to show subtly differentiated expressions. In the following studies, we want to check which features of our approach permit a better recognition of emotions. First, we compare MSE animations to static emotional expressions presented in their apex and MSE animations to animations that do not respect the defined constraints. Second, we look at signals that we have singled out from the sequenced sets of emotional behaviours and we present them one by one. We check if the dynamic animations of short signals that contribute to multimodal expressions of emotions are sufficient per se for a particular emotion attribution.

Hypotheses

Our hypotheses (H) are the following:

H2.1. The recognition rate of multimodal sequential expression is higher than the recognition rate of static displays presented at the apex.
H2.2. The recognition rate of multimodal sequential expression is higher than the recognition rate of single dynamical signals.
H2.3. The animations generated using the constraint-based sequences are more believable than constraintless sequences (i.e., animations not obeying constraints).

Procedure of MSE study 2

This study, accessible through a web browser, was divided into three sections:
S1. Twenty-four static images (24 stimuli)
S2. Sixteen MSE animations presented alone (16 stimuli)
S3. Eight MSE animations presented along with eight “constraintless” animations (16 stimuli)

Images in section S1 show facial expressions, gaze, and/or head movements (neither gestures nor torso movement were used). Since given facial expression have not been specified yet for some emotional states that are used in our evaluation, we have opted for three images presenting three different expressions chosen from the signals that occur in the MSE animations. For each image, we used the keyframe that corresponds to the apex. Each image was shown for 4 seconds (see figure 7).

For sections S2 and S3, the three most dissimilar MSE animations were chosen from a set generated for each emotional state. The choice was based on the presence of different signals and/or their occurrence in time.

Section S2 of the study is composed of animations showing sequences of multimodal behaviours. For each emotion, two different MSE animations were shown. One animation was presented on each web page. The character was not speaking. The duration of each animation was about 10 seconds.

In section S3, 16 animations are presented, eight of which were generated using our algorithm and which satisfied the defined constraints. The other eight present the same nonverbal behaviours, but the order of appearance and duration of each nonverbal signal were chosen manually to be inconsistent with one or more constraints. We call these animations “constraintless” animations.

Participants were asked to recognise the emotions displayed by the virtual character. Each image or video shows the character displaying one emotional state. The dimension of the character head was kept constant in all the images and videos, although a different framing was used (only the head for images and a half body for the animations).

The study was constructed as follows: Each subject has to see all 24 images from section S1 before seeing the animations of the latter sections. In sections S1 and S2, after watching one image or animation the participants have to attribute one emotional label to the perceived emotional state from an 8-elements list before they can pass to another page with a new animation.

Section S3 checks the role of constraints in emotion recognition. On each web page, two animations of the same emotion were presented. Contrary to the other studies, the participants could start, stop, and review the animations. In this section of the study (S3), participants have two different tasks: A recognition task was complemented by a ranking task. First, participants were asked to attribute a label to the depicted emotion, using a similar procedure to the previous sections (S1 and S2). But they were also asked to choose which animation is more believable between constraint-based and constraintless animations.

In all studies, participants were told that they could use each label more than once or not at all.

Results of MSE Study 2

Forty-eight participants took part in sections S1 and S2 of the study (25 women, 22 men, and one gender not stated) with a mean age of 29 years (SD = 7.36), mainly from France (23 %), Poland (21 %), and Italy (12.5 %).

None of them works in the domain of virtual characters. Out of the 48 participants who finished sections S1 and S2, 42 finished section S3 (20 women, 21 men, and one gender not stated) with a mean age of 28 years (SD = 3.98), mainly from France (21 %), Poland (19 %), and Italy (14 %).

Images. In section S1 (still images), a repeated measures ANOVA was calculated to check for the impact of emotions. An effect of emotions on the number of correct recognitions of still images was observed ($F(3.80, 179) = 52.13$, $p < .05$). Overall hit rate interval is very large (4-90 %) while the maximal hit rate varies from 25 (Pride) to 90 percent (Anger). Means, standard deviations, and maximal hit rates for the three presentations of each emotion from section S1 (still images) are shown in Table 17.

Table 17

Static images: mean and maximal hit rate, standard deviation and mean Kappa (k)

Emotion	Mean (%)	Max (%)	SD	k
Anger	70	90	0.22	0.84
Anxiety	24	29	0.04	0.67
Cheerfulness	55	83	0.29	0.72
Embarrassment	38	50	0.14	0.72
Panic fear	46	69	0.36	0,79
Pride	24	25	0.02	0.72
Relief	24	50	0.23	0.73
Tension	30	42	0.11	0.69

A Kappa score was calculated for each image. We modified the calculation of the number of correct rejections, as we did not have an 8x8 design, but eight emotional labels 24 (three images per emotion setting). The lowest value (0.595) was obtained for one Cheerfulness image and the highest value (0.865) for one Anger image (see Table 17 for mean results).

A comparison brought down to labels from the same category was also realised. The lowest k was of 0.179 for an image of Embarrassment in C3 and the highest of 0.761 for an image of Cheerfulness in C2 (C1 - [0.279-0.505], C2 - [0.593-0.761], and C3 - [0.179-0.741]). In C1, the most recognised image was of Relief (= 0.505), followed by one of Cheerfulness (= 0.472). In C2, the most recognised image was that of Cheerfulness (= 0.761), followed by an image of Embarrassment (= 0.688). In C3, the most recognised image was of Anger (= 0.741), followed by Panic Fear (= 0.739).

MSE animations. The hit rate means of the animation presentations from S2 were compared with those from S3.

Linear contrasts showed no difference between the grouped means of presentations one and two (section S2 of the study) and the means of presentation three (section S3) for six out of eight emotions. Only in the case of Embarrassment is the first presentation recognised by 85 percent, while the second and third presentations are similar, with a recognition mean of 42 and 40 %. For Pride, the third presentation is more recognised than the other two.

Consequently, the Presentations displaying the MSE sequence from sections S2 and S3 were considered together.

Repeated measures ANOVA showed an effect of Emotions on the simple hit rate with F(1.56, 30.14.71), p <.05.

The unbiased hit rate shows a nonaleatory attribution for all MSE animations, with the minimum k value of 0.719 for an Anxiety MSE animation and a maximum value of 0.865 for a Panic Fear MSE animation (see Table 18 for mean results).

Finally, a comparison brought down to labels from the C1, C2, and C3 categories was realised for the animation presentations. The lowest k was of 0.289 for an animation of Cheerfulness in C1 and the highest of 0.785 for an animation of Panic Fear in C3 (C1-[0.289-0.590]; C2 - [0.569-0.747]; and C3 - [0.396-0.785]). In C1, the most recognised animation was of Relief (= 0.590), followed by one of Pride (= 0.563). In C2, the most recognised animation was that of Embarrassment (= 0.749), followed by Pride (= 0.736). In C3, the most recognised animation was of Anger (= 0.825), followed by another animation of Anger (= 0.768) and by one of Panic Fear (= 0.768)

Table 18

MSE animations: mean and maximal hit rate, standard deviation and mean Kappa (k)

Emotion	Mean (%)	Max (%)	SD	k
Anger	88	96	0.32	0.85
Anxiety	43	48	0.50	0.72
Cheerfulness	54	64	0.50	0.80
Embarrassment	57	85	0.50	0.76
Panic fear	78	85	0.41	0.86
Pride	65	79	0.48	0.80
Relief	61	79	0.49	0.80
Tension	43	52	0.50	0.75

Effect of constraints. The emotional displays generated with and without the use of the constraints were compared. In section S3, the MSE animation was more often chosen as believable than the constraintless one for the emotions of Anger, Anxiety, Cheerfulness, Panic Fear, Pride, and Relief (as measured with X^2; $p<.05$). Only for Embarrassment and Tension was the choice of the sequential animation not above chance level ($p<.05$).

Comparison of Images and MSE Animations study 2
Section S1 versus Sections S2 and S3
The mean hit rate for images is lower (0.402) than the mean for MSE animations (0.615), and the repeated measures ANOVA shows that the difference is significant ($F(1, 41) = 91.64$, $p<.05$, after a Huyn-Feldt correction).

In section S1, we obtained a very poor recognition of some images. Consequently, when comparing the hit rate between the images and the animations, we relied only on the image and the video with the greatest hit rate for each Emotion. A repeated measures ANOVA calculated on the hit rates of the images and animations that were best recognised showed, after a Huyn-Feldt correction for sphericity, an effect of Emotions ($F(6.96, 284) = 15.14$, $p < .05$) and of Dynamics (i.e., still images versus animations) ($F(1, 41)$, $p = 29.71$, $p < .05$). An interaction effect was also observed for Emotions Dynamics ($F(7, 287) = 5.24$, $p<.05$).

The mean hit rate for the best recognised videos and images is 65 percent. When relying on the attributions to the images and animations with the best recognition proportion per emotion, t-tests were used to compare if the animations are more recognised than images for each emotional state.

The number of participants was of 48 for Anger, Anxiety, Embarrassment, Panic Fear, and Relief, while it was 42 for Fear, Pride and Relief. In the case of Tension, the improvement was not significant, probably due to a small number of participants in the single signals study. For the remaining three emotions, no improvement was observed.

We acknowledge that the improvement of an emotion recognition may be due to the dynamics of the animations or to any of three main features of the MSE algorithm: the different modalities used, the sequentiality, and/or the chosen constraints. The lack of improvement for some states could simply be due to an insufficient number of annotated videos and not to the insufficiency of the model. In some cases, the recognition is already very high for some still images and single signals, as for Cheerfulness or Anger.

This could be due to the fact that the expression relies on a key signal, sufficient for the recognition of a given state. Indeed, in the single signals study, they obtained a high result for (96 and 70 %). For others (e.g., Tension), the expressive qualities of a behaviour may be guided mostly by expressivity characteristics, while, in our animations, these were kept constant.

In particular, in the case of Cheerfulness, adding supplementary information may not disambiguate the expression but may drive the attributional process away, increasing the chance of attributing alternatives. Indeed, MSE animations of Cheerfulness had a very low score in C1 and much higher for still images. For Anger, the recognition was already very high in the still images (90 % correct recognitions) and the improvement was not significant (96 %).

In the case of Anxiety and Tension, the improvement due to MSE may be less marked as annotations seem to show that this state may be particularly expressed by cues that are presented individually at longer periods of time.

Moreover, the expressive qualities of a behaviour are very important, while the expressivity of the character was constant through the animations.

The presented results show that dynamical and multimodal expressions generated with our algorithm enable our character to communicate many emotional states more efficiently than through static facial expressions or dynamical single modality expressions. We have also found that constraints play an important role in multimodal sequential expressions. Indeed, in most of the cases, the MSE were considered more believable than the animations not respecting any constraints (six out of eight cases).

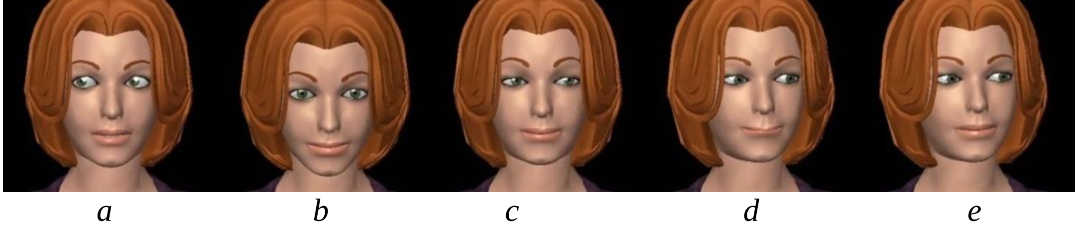
MSE discussion

Our first challenge was to evaluate the MSE algorithm, without an a priori defined lexicon and without any former verification of particular constraints. We acknowledge fully that the results are dependent on the quality and quantity of processed and integrated information, whether from literature or from annotations. Indeed, some emotions like tension had relatively worse scores than the other emotions (e.g., panic fear or relief). This may show that even though our approach is efficient for certain emotions, it may not be sufficient to generate adequate displays for all emotional states. For instance, in the case of tension and anxiety, such cues as the expressive quality of the behaviour are particularly important.

Another limitation of this study is the use of forced choice, with the possibility of choosing only one label out of a set. This procedure is often used in the perception studies on emotional displays (among others by Ekman, Friesen, 1975), but it may force the user's interpretation of the expressive behaviours (see Haidt, Keltner, 1999). We also use only one virtual character's face in this study, while studies show that the interpretation of generated behaviour may be influenced by the character's physical characteristics, such as prominence of eyebrows or character's gender (Hess, Adams, Kleck & 2004). What is more, due to the limitations of the Greta character, we could not exploit all the possibilities of multimodal communication. For example, we did not use posture to express emotional states, which is an important channel to communicate emotions (Kleinsmith & Bianchi-Berthouze, 2007).

We acknowledge that this work was placed in a context-free setting, where emotional displays are to be recognised only from visual cues. The communication of subtle emotional states probably cannot be fully successful without considering situational context (e.g. nonverbal behaviour of the receiver). It would have been interesting to see differences in attributions in a well defined context, for example by placing the virtual character in an airport, in interaction with a hostess, or in any other well defined scenario as will be described in the next study. Many other factors such as, e.g., position of the sender and receiver, available modalities, or other communicative intentions to be communicated at the same time, etc., may also influence multimodal affective behaviours. These limitations may explain why some recognition rates are still somehow low.

In figure 3 an example of the animation for the expression of embarrassment is shown. The images present the frames of the animation of Greta displaying respectively the signals: a) look_right, b) head_down and gaze_down, c) gaze_left, d) gaze_left and non-Duchenne_smile, e) gaze_left.



a *b* *c* *d* *e*

Figure 3. *An example of an expression of embarrassment.*

Figure 4 shows another example of animation of the same emotion. In Figure 4 the following signals are displayed: a) neutral expression, b) smile, c) smile and gaze_right, d) gaze_left, e) gaze_down and head_down, f) touching face gesture.

a *b* *c* *d* *e* *f*

Figure 4. *An example of multimodal expression of embarrassment*.

We stress, however, that besides creating well recognised expressions, the development of MSE has allowed a greater malleability of expression to the Greta character. In consequence, the character became a more flexible tool for experimental settings. It allowed the reproduction of human behaviour observed in a naturalistic setting, during a passenger's face to face interaction with an airport hostess.

4.3. Research question

Given that correlational studies on naturalistic human behaviours do not allow determining the effect of specific cues on what inner states are attributed by third party observers, an idea is to generate behaviours using a virtual character. The presence/absence of what nonverbal cues lead humans to attribute appraisal and emotional labels when they observe strangers?

In face to face interactions, behaviours that we observe in others are a rich source of information on the person's state of mind and his/her reaction to the environment and their decoding may be vital for everyday life. Which cues are sufficient per se to increase some attributions and the removal of which "creates an imbalance" and increases the ambiguity of an attribution, decreasing the perception of some of the linked emotions or appraisals?

While our *Naturalistic expression study* (described in detail in *Problematics 1: Appraisal components of emotion*) consisted of linking nonverbal behaviour performed by targets to attributions of appraisal and emotion by third parties, this study, *Manipulated expression study* relies on synthetic expressions generated with a virtual character to verify this link between present or absent cues in complex expressions and attributions.
Laypeople were asked to participate in the study and evaluate short video clips (7-23 sec long) by using likert scales to attribute emotions and appraisal checks to targets.
Greta, the virtual character, was used to generate complex expression and to test the impact of observed cues.

4.4. Predictions

Five AU were singled out as of significance, in relation to theory and based on human study correlations. These were AU 7 (positive correlation with coping observed in *Naturalistic* *expression study*), AU 4 (positive correlation with obstruction confirmed in *Naturalistic expression study* and also expected with on Anger inliterature), AU 12 (negative correlation with obstruction confirmed in *Naturalistic expression study*), AU 2 (positive correlation with relevance/discrepancy observed in the Naturalistic expression pre study) and AU 5 (positive correlation with Fear confirmed in *Naturalistic expression study*).

AU 12 and AU 5 have not been kept as cues to be manipulated in this study as they did not appear sufficiently in the selected videos.

Thus, four conditions have been created: the full reproduction of the set, that is the control condition (cond100) and each set with the removal of the presence of AU 4 (cond4), of AU 7 (cond7) and of AU 2 (cond2).

4.4.1. Theoretical hypotheses

The removal of a particular action unit from a sequence of behaviours is expected to have an impact on the emotion and appraisal scales associated to that action unit.

4.4.2. Operational hypotheses

We have six operational hypotheses (H).

H1. We expected a difference between the conditions of presented clips for several emotion and appraisal scales:

- H1.1 Exterior part of brow raise (AU 2) on relevance/discrepancy
  - i) Removing AU 2 will decrease attribution of discrepancy
- H1.2 Frown/brow contraction (AU 4) on obstruction
  - i) Removing AU 4 will decrease attribution of obstruction
- H1.3 AU 4 on Anger
  - i) Removing AU 4 will decrease Anger
- H1.4 Eyelid contraction (AU 7) on coping
  - i) Removing AU 7 will decrease attribution of coping

H2. We expect no condition x emotional scale interaction

H3. We expected no general impact of condition on the attributions when all scales are considered.

H4. We expect an impact of emotion scales on attributions, as we expect more negative emotions to be attributed.

H5. We expect an interaction between video and particular scales

H6. We expect no interaction effect of condition x video.

4.5. Analyses

To observe if the particular components of general expressions presented in a set of coordinated behaviours have an impact on the associated emotion and appraisal scales different conditions of presentations were compared using analysis of variance (ANOVA). Repeated measures ANOVAs are performed to check if the condition (an AU being present or absent from a general behavioural expression) has an effect on the attribution of a specific emotion/appraisal score to judged videos.

4.6. Materials

4.6.1. Stimuli corpus creation

"Target" videos

Videos have been drawn to be reproduced by a virtual character. The selection was controlled- aleatory, with a control of gender (two male targets, two female ones). Mean duration of selected 4 videos is of 12 sec (min = 7, max 23):

Table 19

Description of Target Videos

Clip	Gender	Duration
45a	F	23
60c	F	9
70b	M	7
110c	M	9

Four different profiles of virtual characters (2 males, 2 females) have been used to reproduce the sets of behaviours. The same facial features were kept for the two profiles per gender, and the identity was manipulated solely through haircuts, presence of glasses and skin texture and colour.

Illustration 1 ***Four virtual character profiles in neutral expression.***

4.7. Procedure

4.7.1. Set-up

The evaluation involving a virtual character was run online. Each participant accessed the study individually through a web browser, having received a link by email.

Participants were told in the invitation mail to prepare headphones before starting the study, as these were needed for spoken guidelines.

4.7.2. Questionnaire presentation

After a first page gathering basic demographic information on participants, general guidelines and aims of the first study were presented by a female virtual character. The purpose of this general guideline was to habituate the participants to the general behaviour of virtual characters and their limitations (e.g. in their bodily and voice expression). The facial model for the virtual character used in the introduction is different from the four facial models used in the actual study.

The character speaks out the guidelines (using MARY text to speech, see Schröder, Pammi & Türk, 2009; female voice) and uses basic pointing gestures and head nods. The character spoken guidelines are followed by a more detailed guideline in traditional text format (Appendix VII).

For the study, participants view silent videos and answer a set of questions after each one.
A pre test showed that reproducing with the virtual character the same paradigm as for the *Naturalistic expression study* overwhelms the participants, who complained about the length of questionnaires. In consequence we opted for presenting one appraisal question drawn randomly out of the pool of 7 per video. Thus, each evaluation set consists of 8 likert scales: one appraisal attribution and seven attributions of emotion intensities (see Appendix VII).

As in the *Naturalistic expression study*, sentences evaluate suddenness, goal obstruction, detection of an important and unexpected event, coping potential, internal and external standards violations. Participants answer on a 7-point likert scale (from 0 = totally disagree to 6 = totally agree). With such a measure an average significantly above 3 can be considered to confirm the presence of an appraisal and an average significantly below to confirm the absence of an appraisal in a given clip.

The emotion scales following the appraisal question ask if the observed passenger is experiencing Joy, Anger, Relief, Sadness, Contempt, Fear and Shame. Two orders of presentation were used. Each emotion is evaluated on a separate likert scale from zero (no emotion) to six (strong emotion). An emotion can be considered attributed when the average is significantly >0.

The study was created in the PHP language and can be accessed through a simple web browser. The clip-stimuli were presented on full screen (in the original size).

Participants were randomly attributed to one of the 4 full length clip sets.

4.8. Results

4.8.1. Descriptive statistics

118 participants took part in the study (40 women, 77 men and 1 undeclared). The mean age of participants was of 31 years (SD 11.25). The majority of participants reported they have spent most of their life in France (70 %), followed by Tunisia (6 %), Italy (6 %), Morocco (3 %) and Germany (2 %). Other countries were cited by less than one percent each (Argentina, Belgium, Burkina Faso, Chile, China, Colombia, Germany, Ivory Coast, Uganda, Romania, Spain, Sweden, Switzerland, Vietnam).

The majority of participants (42 %) had a postgraduate level of education and above (PhD, MAS, MBA), 24 % had a master degree, 11 % a bachelor degree, 21 % had started university/college and 2 % stopped with a high school level.

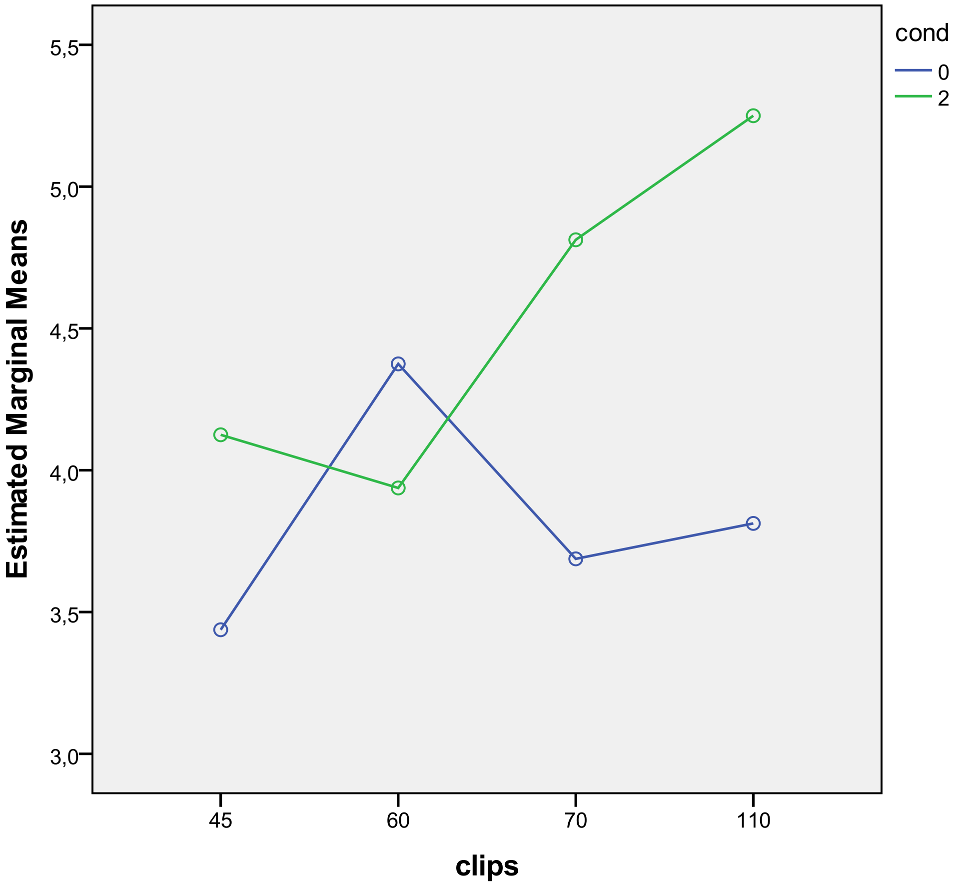
The majority of participants came from the computer science/engineering domain (58 %), then maths/statistics (12 %), psychology and other social sciences (7 %), economics/finance (5 %), medical studies (1 %) and other (17 %).

4.8.2. Impact of behaviour presentation manipulations

R1. Specific results: comparing conditions per emotion

R1.1 AU 2 on relevance

A repeated measures ANOVA on attributions of relevance/discrepancy appraisal showed an effect of condition 0 versus 2, with $F(1, 15) = 4.741$, $MSE = 3.337$, $p<.05$.

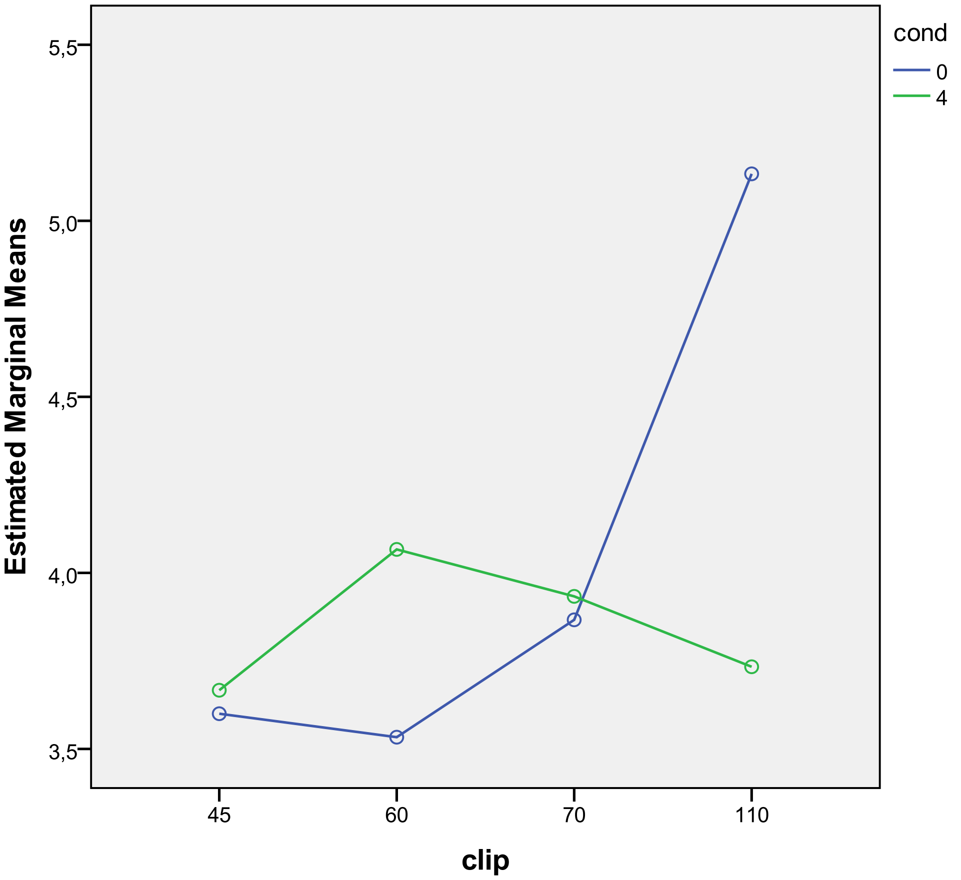


Graph 1. *Mean of relevance/discrepancy attributions to the fully reproduced behaviours (condition 0) for four clips and reproduction of these same clips' behaviours without any raise of outer part of eyebrow (condition 2)*

As we see on graph 1, in three out of four clips, there is a trend that shows that removing an outer eyebrow raise (AU 2) leads to an increase in the attribution of relevance/discrepancy.

R1.2 AU 4 on obstruction

A repeated measures ANOVA was run on appraisal of obstruction attributions to 4 video clips and showed no effect of conditions 0 versus 4, with $F(1, 14) = .279$, $MSE = 3.615$, $p>.05$.

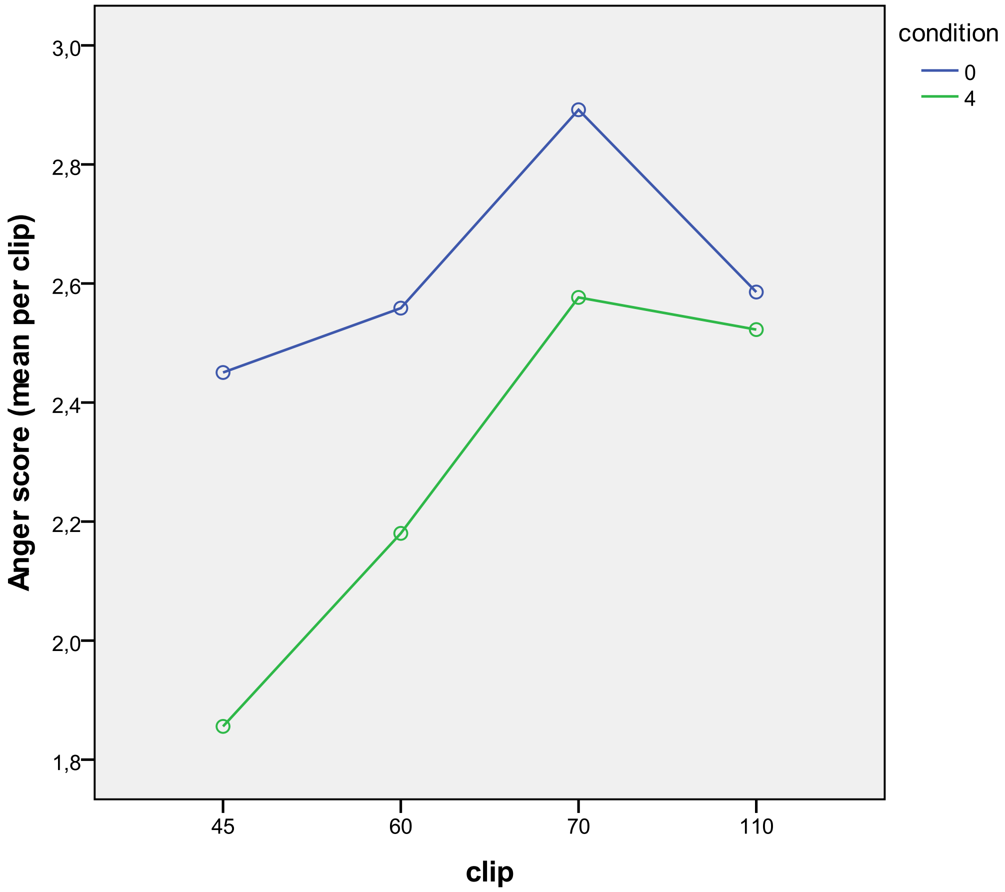


Graph 2. *Mean of goal obstruction attributions to the fully reproduced behaviours (condition 0) for four clips and reproduction of these same clips' behaviours without any frown (condition 4)*

As we see on graph 2, removing a frown/eyebrow contraction (AU 4) does not lead to a clear tendency (either in increase or decrease) in the attribution of goal obstruction.

R1.3 AU 4 and Anger

A repeated measures ANOVA on attributions of Anger showed an effect of condition 0 versus 4 on attribution of Anger, with $F(1, 110) = 11$, $MSE = 2.306$, $p<.001$.

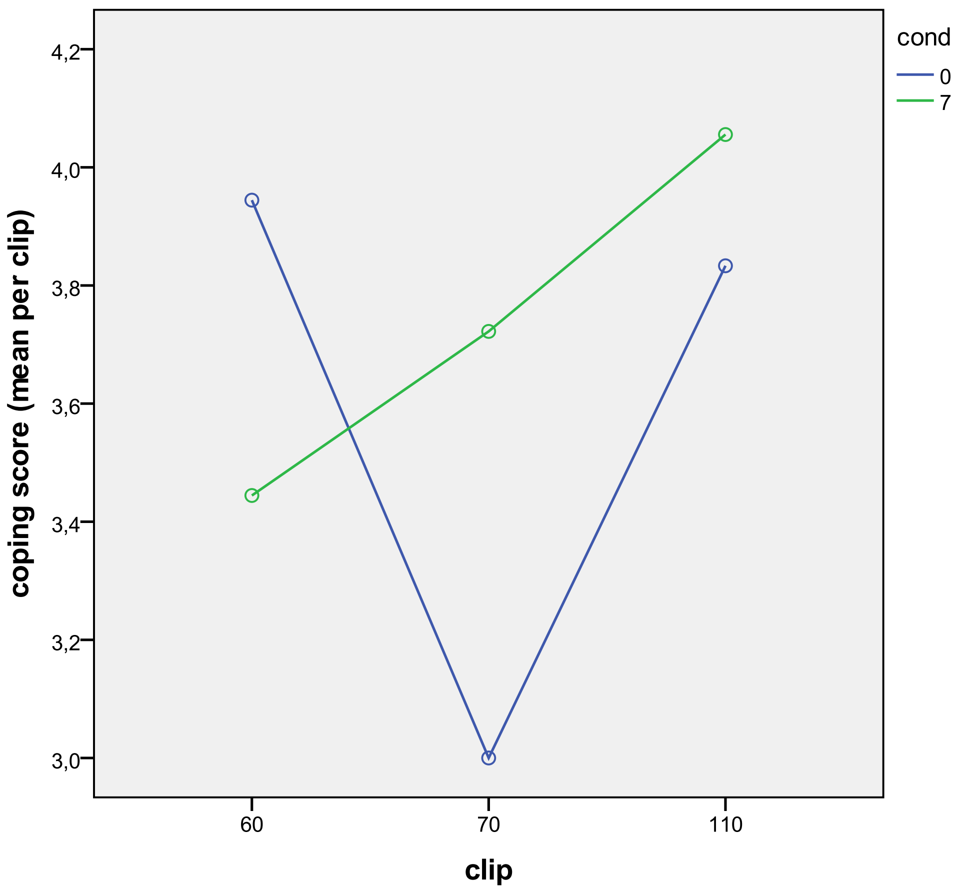


Graph 3. *Mean of Anger attributions to the fully reproduced behaviours (condition 0) for four clips and reproduction of these same clips' behaviours without any frown (condition 4)*

As we see on graph 3, removing the frown (AU 4) leads to a decrease in the attribution of Anger.

R1.1 AU 7 on coping

A repeated measures ANOVA was run on the three clips containing condition 7 (clip x condition). There was no effect of the condition 0 versus 7 on attribution of coping, $F(1, 17) = .217$, $MSE = 2.730$, $p>.05$.



Graph 4. *Mean of coping attributions to the fully reproduced behaviours (condition 0) for three clips and reproduction of these same clips' behaviours without any eyelid contraction (condition 7)*

As we see on graph 4, removing an eyelid contraction (AU 7) does not lead to a clear tendency (either in increase or decrease) in the attribution of coping potential.

General results

Separate statistical analyses were run on emotion and appraisal scales, but results are provided together.

A repeated measures ANOVA was run with a 4 videos x 3 conditions (100, 4, 2) x 7 emotions (Joy, Anger, Relief, Contempt, Sadness, Shame, Fear) design.
and a 3 videos (60, 70, 110) x 4 conditions (100, 7, 4, 2) x 7 emotions (Joy, Anger, Relief, Contempt, Sadness, Shame, Fear) design.

For appraisals, a repeated measures ANOVA was run with a 6 appraisals (suddenness, obstruction, discrepancy, coping, respect of internal standards, external standards violation) x 4 video (45, 60, 70, 110) x 3 conditions (100, 4, 2) design.
For Emotion a similar design was used with 4 videos (45, 60, 70, 110) x 3 conditions (100, 4, 2) x 6 appraisals (suddenness, obstruction, discrepancy, coping, respect of internal standards, external standards violation) set up.

R2. No general interaction effect of **Emotion x conditions** on attributions on four videos with 3 conditions ($p>.05$) nor on three videos with 4 conditions ($p>.05$). The same case was for **appraisal x condition** interaction, which was non-significant ($p>.05$ for both).

R3. When looking at emotion attributions for three videos that have all conditions, no effect of **condition** was found, with $p>.05$, nor in four videos with three conditions with $p>.05$. Similarly, when looking at appraisal attributions no effect of condition was found either in the 3 clips x 4 conditions, nor in 4 clips x 3 conditions ($p>.05$).

R4. When looking at attributions to four videos with three conditions, multivariate ANOVA showed an effect of the **Emotion** scales with $F(3.95, 430) = 122$, $MSE = 3.31$, $p<.001$ and of the **Appraisal** scales with $F(3.47, 45.2) = 4.587$, $MSE = 6.19$, $p<.05$.
This effect was also significant in attributions to three videos with all four conditions when looking at Emotion scores with $F(3.77, 410) = 111$, $MSE = 3.58$ $p<.001$ and at Appraisal scores with $F(3.29, 42.89) = 5.04$, $MSE = 6.50$, $p<.005$.

R5. We expect an interaction between video and particular scales.
An interaction between **Emotion and videos** was found ($F(7.49, 816) = 24.62$, $MSE = 3.38$ $p<.001$ for 3 videos; $F(10.31, 1124) = 18$, $MSE = 3.53$, $p <.001$ for 4 videos)

R6. For four videos with three conditions and for three videos with four conditions there was no significant **condition and video** interaction for appraisal ($p>.05$ for both) nor for Emotion ($p>.05$ for both).

4.9. Discussion

In this chapter, we present a novel methodology for testing the impact of particular cues on emotion and appraisal perception in others. While relying on spontaneous human behaviour, we decided to observe the contribution of specific facial action units in the complex sequence of behaviours by removing them from the holistic picture, one unit at a time. We decided to use a virtual character, Greta, to display the nonverbal behaviours seen in the Lost Luggage naturalistic corpus (Scherer & Ceschi, 1997) described in former chapters. Greta has the possibility to generate face expressions based on FACS, as well as torso and other body movements. However, to enable the display of face and body movements as in the chosen emotional extracts from the corpus, the capabilities of Greta character had to be developed to include presentations in sequences of units, with different starting and ending times for particular face and body cues, and superposition of different units.

Thus, we created a novel approach to the generation of emotional displays in a virtual character, leading to the creation of the Multimodal Sequential Expression (MSE) model. The emotional expressions generated with this model can not only be displayed on the face, but also as body cues, using different modalities as well as signals that occur in sequences. This approach allows for high flexibility and variation of emotional displays, which is both necessary for the use of Great character in our perceptive study as well as possibly useful to increase the range and specificity of emotional displays of virtual characters.

Once the behaviour generation capacities of Greta were modified, evaluation studies were run to verify three main features of our approach, which are multimodality, sequentiality, and the use of constraints. The results of our first study show that the recognition of the MSE animations is high. The second study enabled us to further observe that multimodal sequential expressions are better recognised than static emotional displays in their apex and (at least for some emotional states) better than dynamical single signals. It also showed that the application of constraints increased the believability of the multimodal sequential expressions. Once those MSE evaluations were performed and their results satisfying, Greta was used to reproduce FACS and body behaviours observed in the human corpus extracts. A perceptive study on the link between emotions, appraisal and nonverbal behaviour was run, with a similar design to the one presented in chapter *Problematics 1: Appraisal components of emotion.*

In the *Manipulated expression study*, results confirm the majority of our expectations. First, our manipulation of particular nonverbal cues does have an effect on some of the predicted emotion and appraisal attribution scales. Also we do not see a general condition and scales interaction, which confirms our predictions. Thus we could argue that changing one cue does not obligatorily lead to a change in the holistic perception of all scales, e.g. through interactions with other present behaviour, but does affect at least some of the expected appraisal dimensions. Thus, we see an effect of AU 2 (outer brow raise) on relevance/discrepancy and of AU 4 (frown) on Anger.

The co-presence of AU 7 (lower eyelid contraction leading to eye aperture narrowing) with attribution of coping that was observed in *Problematics 2: Expressions of appraisal and emotion,* was not found. The removal of the lower eyelid contraction from character displays did not decrease coping attributions.

Similarly, the co-presence of AU 4 (frown) with higher obstruction attributions was not observed in this study, as our manipulation of removing the frown from displays did not produce any significant differences on attributions on that scale.

Moreover, as predicted, when all scales are considered, no effect of the condition was also shown.

Our results also show that scales matter both per se and in interaction with videos. Thus, scores were not equally distributed across scales – some appraisals and emotions were given generally higher scores than some others when all videos were considered. Also, scores on scales were different from video to video.

No interaction effect of condition and video was expected, all scales considered, and none was found.

4.10. Conclusion

Thus, the way facial expressions were decoded suggests that particular cues do have an impact on attribution of related appraisal dimensions, without obligatorily having an impact on all other scales. So what kind of model do these results support?

Smith and Scott (1997) contrasted three models, which they applied to the encoding process: the purely categorical, the componential and the purely componential models. What they called the *purely categorical model* consists of conceptualising emotional expressions as basic interpretable units. The components of an expression, like an eyebrow movement or a wide opening of the eyes, are considered meaningless. Similitudes between expressions, such as for instance two expressions sharing the same facial actions, do not provide interesting information in terms of associated mental states. The occurrence of wide eyes opening in surprise and fear expressions would be considered a product of chance rather than bearing some function. As a consequence, the specific facial actions contributing to an emotional expression do not provide any help to interpret the emotion.

The second model the authors present is the *Componential model*, which underlines the contribution of particular components of facial expressions. In this model, at least some of the individual components, such as for example individual action units, are considered meaningful per se, contributing directly to the general meaning of an expression. In that stance, expressions are seen as having a *systematic, coherent, and meaningful structure* (Smith & Scott, 1997). The holistic expression conveys more information than what is provided by the sum of the messages conveyed by the individual facial actions contributing to it.

A third model, the *Purely componential model*, states that a holistic facial expression does not convey anything more than the particular facial actions that contribute to it. The full configuration is nothing more than the sum of its components and the interactions between components do not have any effect.

In our study we explored the contribution of particular facial action units to the meaning attributed by participants to the general expressions and behaviours they see. We show that altering one component, in our study the removal of an action unit from a sequence, does impact judgment. We argue that our results contradict the first model, the purely categorical

one, given that the manipulation of one cue can at least have a direct impact on an associated appraisal dimension.
The removal of an action unit never leads to a total rejection of associated emotion and appraisal labels, although it alters the scores. It can be seen as an indication that individual action units that we looked at are not necessary for the attribution of particular states.
We will not generalise this finding to other, non explored, behaviours as some could be essential in some mental state expressions, the smile being described for instance in literature as a key expressive cue in joy (e.g. Ekman & Friesen, 1982).

What is more, although our study supports the componential approach, it does not allow us to differentiate between the componential and the purely componential approaches. Other studies show however that the full facial expression has a different effect on attribution than the sum of particular facial actions. Some studies have explored the meaning attributed to one particular facial action presented one at a time and compared it to the judgment of combinations of several actions displayed in one composite expression (e.g. Bevacqua, Heylen, Pelachaud & Tellier, 2007). Results show that meanings of combinations of actions may not correspond to the sum of meanings attached to each of the action component.
What is more sequence and timing of expressions have been described as important in the decoding and encoding of internal (e.g. With & Kaiser, 2011; Niewiadomski, Hyniewska & Pelachaud, 2011). This emphasis on the impact of sequence of components of expressions contradicts the idea of simply additive effects of present behaviours.

5. General discussion

...it seems to me that if I were to become corporeally anaesthetic, I should be excluded from the life of the affections, harsh and tender alike, and drag out an existence of merely cognitive or intellectual form. Such an existence, although it seems to have been the ideal of ancient sages, is too apathetic to be keenly sought after by those born after the revival of the worship of sensibility, a few generations ago.

James, 1890

This PhD thesis focuses on how laypeople perceive another person's mental states. We explore the interaction between appraisals (cognitive evaluation of a situation by a person), emotions and the associated nonverbal behaviours. To achieve this goal, we conducted two perceptive studies, the first one on naturalistic behaviour and the second one reproducing some of this behaviour through a virtual character. Both studies involved participants asked to watch videos and assess the behaviours they observe. We summarize here the main results and findings of the two studies. Then, we discuss the limitation of the presented approach, and trace some future research lines emerging from our study.

5.1. Study 1: Relation between internal states and behaviours in humans

In the first study, we explore the link between emotions and appraisals in a new context, which has not been formerly explored. We base our exploration on an audio-visual corpus filmed by a hidden camera, to study nuanced and complex behaviours in a naturalistic setting. The use of covert recordings in everyday emotional situations has a considerable advantage over acted portrayals. It enables to go beyond the prototypicality of some expressions, and eventually overcome bias created by folk-psychological beliefs about the association between emotions and facial displays (Reisenzein, 2000a; Russell, Bachorowski & Fernández-Dols, 2003). The videos in the adopted audio-visual corpus present two persons interacting in a spontaneous manner, in a situation of a passenger claiming the loss of a luggage to a hostess. The participants (N = 122) were shown videos of passengers and were asked to decode their behaviours. We carry out this study with two purposes in mind. The first one is determining the relation between emotions and appraisals, by assessing how laypeople attribute internal states, namely emotions and appraisals, to individuals they observe in audio-visual recordings. We analyse the nature of these relations based on the attribution scores of participants. The second purpose is to analyse how participants judge spontaneous expressions and how they attribute internal states based on this observation. Participants evaluate the presence of an internal state by assigning a score on a likert scale. These scores are then linked to the annotations of spontaneous behaviours previously carried out by experts, i.e., professionals in the field of nonverbal expressions.

This enables to analyse whether the presence of some appraisals is necessary and/or sufficient for the attribution of some emotions.

The results of this first analysis only enable to show the contingency of particular appraisals and emotions, without providing any information on causality. According to necessity and sufficiency indices, no single appraisal attribution is sufficient to predict whether or not an emotion will be attributed. The attribution of an appraisal like suddenness, for example, may but will not necessarily bring about the attribution of an emotion, like fear, anger, etc. Thus, no emotion is characterised by one appraisal alone.

Results also show that no appraisal is strictly necessary for the attribution of any emotion. A particular appraisal may be associated to an emotion, but if other elements contributing to the same emotion are present in the observed behaviour, the observer may not perceive a particular appraisal and yet still attribute the emotion. Thus, we conclude that there are no one-to-one mapping rules between appraisal patterns and reported emotion labels (following Kuppens, 2010).

These results are consistent with literature, as it is expected that emotions are predicted better by combinations of appraisals than by individual appraisals (Roseman, 1984).

We also explore the relationship between nonverbal behaviours and internal states of the persons in the videos, as observed by the participants. To this end, we annotated face and body actions. For the face, we used the Facial Action Coding System (Ekman, Friesen & Hager, 2002). For the body, we created a scheme appropriate for the specific situation of claiming a loss of luggage. In each video extract, we measured the sum of the duration of each facial expression and of each particular body movement. We also calculated the mean duration of each action and we noted the frequency of that action in each extract. We then perform correlation analyses between each of these three measures of behaviour and the judgments reported by participants by scores attributed on likert scales.

Participants' attributions show to a certain extent the associations already reported in the literature.

For appraisals, we found a positive correlation between coping potential and lower eyelid contraction and raised chin with pressed lips. For goal obstruction, only a negative relation was found with the smile – no other predictions were confirmed. Suddenness, relevance, discrepancy and the two norm standard appraisals predictions were not confirmed.

For emotions, the majority of associations with behaviours confirm the hypotheses we made based on the descriptions reported in the literature.

A negative correlation was found between lower eyelid contraction and Anger, but no other expectations were confirmed.

For Fear, a positive correlation was found with the opening of the eyes aperture and a negative one with lower eyelid contraction.

For Joy, a positive correlation with smile, inner brow raise and outer brow raise was apparent, but no such positive correlation was found with cheek raise, nor whole brow raise (inner and outer at the same time) or other predictions.

For Sadness, there was a positive correlation with inner brow raise and opening of the mouth.

In general, our results confirm that a correlation exists between specific patterns of behaviours and mental state attributions. We remark that correlations demonstrate only a contingency between the behaviours and emotional states observable in the presented corpus. However, no kind of causal relation can be inferred.

Also notice that the use of correlations gives rise to attributions biased by all kinds of interrelations between face and body actions, as some of these actions naturally occur more often with some others.

5.2. Study 2: Relation between internal states and behaviours in a virtual character

In the second study, we used a virtual character, Greta (Bevacqua et al., 2010), as a novel tool to experimentally study the influence of specific facial actions on the perception of emotions and appraisals by observers. We presented one facial action at a time and displayed it along with various other facial and body behaviours in complex sequences. This process enabled us to check the impact of removing single facial actions from the holistic picture.

Greta character's behavioural capacities had to be improved in order to enable better synchrony both between the character's consecutive actions and the superposition of different actions. We thereby created a new way of generating expressions for characters, the Multimodal Sequential Expression (MSE) model. In a first study, we showed that the recognition of the emotions displayed by the Greta character and generated according to the MSE model is high. The second MSE study showed that multimodal sequential expressions are better recognised than static emotions displayed in their apex and better than dynamical cues presented alone rather than in a sequence. It also showed that respecting an order of display, specific to the presented sequence of behaviours, increased the believability of expressions.

Greta was then used to generate the behaviours that had been observed and annotated in the Lost Luggage corpus extracts. A perceptive study on the link between emotions, appraisal and nonverbal behaviour was conducted. Participants (N = 118) were shown behaviours displayed on four embodiments of virtual characters (two male faces and two female faces) and were asked to evaluate the behaviours using emotion and appraisal scales. The great majority of our expectations were confirmed.

From the results we collected, we can infer that removing one facial action from a sequence of behaviours does have an effect on some associated appraisal or emotion, without necessarily changing the holistic perception of the entire sequence of behaviours. The results show that from a general point of view, the removal of single facial actions did not have a global effect on the perception of entire sequences nor on the perception of the displayed videos.

We can argue that our results go against the purely categorical approach, as our manipulation does contribute to change attribution scores of participants. We show that the removal of one cue can have a direct impact on an associated appraisal dimension.

Following Smith and Scott (1997) and based on our results, we are inclined to support what the authors called the Componential model. The model underlines the contribution of particular components of facial expressions, although the holistic expression could eventually convey more than information provided by the sum of the messages from individual facial actions contributing to it. However our study does not allow us to differentiate between componential and purely componential perception of expression - we cannot say the full facial expression has a different effect on attribution than the sum of particular facial actions. It would still be necessary to test the impact of particular action units. What is more, we have to keep in mind the now well highlighted occurrence of sequences in emotions and emotional behaviours. On the emotion side, Scherer's Componential Process Model (CPM; Scherer, 2009) emphasises the *dynamic unfolding of emotion*, with the sequence of appraisal checks and associated synchronised responses. On the encoding side, With and Kaiser's (2011) work shows the occurrence of facial action sequences specific to recalled emotional states during a sharing task. Keltner and his colleagues have also showed that some sequences are specific to particular emotional expressions (Keltner, 1995; Shiota, Campos, Keltner, 2003).

5.3. Discussion and limitations of the current approach

In the present study on how individuals perceive mental states in others, we are limited by the audio-visual corpus we use. We observe one specific situation, which is that of a loss of luggage at an airport, thus the mental states observed and associated emotions and appraisals are quite restricted by the negativity of the scene. The emotions and appraisals present are a sample which is not necessarily representative of everyday states. Besides, the relationship between appraisals and emotion that we observe here is not to be seen as one to be generalised –the observed relations may be biased by the characteristics of our specific situation. What is more, in those videos the focus is on face to face interactions in which passengers ask a hostess for a service, which probably implies some strategic expression management to have a positive impact on the hostess.

Thus, although videos come from a hidden camera, they probably show individuals using some social self-presentation strategies, inhibiting or masking some expressions which may not be of an advantage in their request. In other words, the recorded displays of authentic expressions of individuals in an emotional situation do not allow to claim that these are pure expressions of emotions, unaffected by push and pull effects (e.g. Scherer, 1985). Observed behaviours are likely to be under the influence not only of display rules but also of emotion regulation (e.g. Scherer & Ellgring, 2007).

We do not claim that emotions should be explored in non-social contexts. On the contrary, social contexts are adequate for the study of emotions, as a number of emotions arise in such contexts. Our setting is ecologically valid and does show various persons reacting differently to a similar event. However, it is necessary to explore a panel of various situations which would be categorised in terms of probable appraisals that could be triggered by the occurring events.

5.4. Contributions

In the first part of this thesis, we explore naturalistic face-to-face behaviours from a hidden camera, that we annotate using FACS (Ekman, Friesen & Hager, 2002) for the face and an internally created scheme for body annotation. We show that in third party judgment of behaviours, no appraisal attribution seemed necessary nor sufficient for the attribution of any emotion label. However, correlation analyses confirm expectations concerning the two concepts, mostly as defined in the CPM model (e.g. Scherer, 1988; Scherer & Ellgring, 2007). We also show the co-occurrence of different facial and body behaviours with specific emotions and appraisals.

In the second part of the thesis, we use a novel method to explore face expressions and the specific impact of particular behavioural cues on emotion and appraisal perception by third party observers. Greta, the virtual character, can display expressions and enables the control of behavioural cues. We contribute to the development of Greta's nonverbal behaviour, by introducing multimodal sequential expression (MSE) displays capacities. These capacities are used to reproduce entire sets of expressions observed in the videos of human behaviour. Our results show that the removal of one behaviour from a set of cues contributing to a general expression, does have a significant impact on the associated emotion and appraisal scales.

Our results therefore contradict the purely categorical model described by Smith and Scott (1997), whereby each specific component of an expression is considered meaningless.

5.5. Future research directions

Nonverbal behaviour goes beyond face and body movements that we have analysed in this PhD thesis. When running perceptive studies we opted for the most ecologically valid presentations of behaviours, which also implied not muting the audio-visual recordings. In consequence participants saw movements and posture changes, but also heard the interaction held between passengers and the hostess. Thus, when evaluating videos of passengers, participants were influenced by verbal as well as non-verbal cues, including vocal changes that are also considered rich indicators of internal states (Bänziger, Scherer, 2005; Grandjean et al., 2005; Laukka, 2005. The text spoken by passengers was transcribed and no emotional labels were present. Looking at the impact of vocal patterns is an essential step to the contribution of presented work.

The direct continuation of started behavioural cue manipulations using Greta character is also planned. We plan to study the impact of particular actions contributing to general expressions, one cue at a time, presented in the same "virtual passenger" context and without a context. This will enable to see if the meaning attributed by laypeople to composite expressions go beyond the sum of particular actions contributing to it.
Studies will also be run with the idea that individual action units are not sufficient for the attribution of an emotion or appraisal. Impact of sequence and timing (speed, duration) of actions will be investigated.

Bibliography

Abrilian, S., Devillers, L. Buisine, S., & Martin, J.-C. (2005). EmoTV1: Annotation of Real-Life Emotions for the Specification of Multimodal Affective Interfaces. In: *Proceedings of the 11th International Conference on Human-Computer Interaction*. Las Vegas, NV.

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience, 4*(3), 165-178.

Adolphs, R. (2004). Emotional vision. *Nature Neuroscience, 7*(11), 1167-1168.

Albrecht, I., Schröder, M., Haber, J., & Seidel, H. (2005). Mixed feelings: expression of non-basic emotions in a muscle-based talking head. *Virtual Reality, 8* (4), 201-212.

Alvarado, N., & Jameson, K.A. (2002). Varieties of anger: The relation between emotion terms and components of anger expressions. *Motivation and emotion*, *26*(2), 153-182.

Arnold, M.B. (1960). *Emotion and personality: Vol. 1. Psychological aspects*. New York: Columbia University Press.

Aviezer, H., Hassin, R.R., Ryan, J., Grady, C., Susskind, J., Anderson, A., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, 19, 724–732.

Adler, R.S., Rosen, B., & Silverstein, E.M. (1998). Emotions in negotiation: How to manage fear and anger. *Negotiation Journal*, *14,* 161-179.

André, E., Rist, T., van Mulken, S., Klesen, M., & Baldes, S. (2000). The Automated design of believable dialogues for animated presentation teams. In: Cassell et al. (Eds.): *Embodied conversational agents.* Cambridge: MIT Press.

Apostle, H.G. (1981). *Aristotle's on the soul (De Anima)*. Grinnell, Iowa: The Peripatetic Press.

Aristotle (1954). *Rhetoric* (Roberts, W. R., Bywater, I., & Solmsen, F., Trans). New York: Modern Library.

Arya, A., Di Paola, S., & Parush, A. (2009). Perceptually valid facial expressions for character-based applications. *International Journal of Computer Games Technology*.

Banse, R., & Scherer, K.R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*, 614–636.

Bachorowski, J., & Owren, M.J. (1995). Vocal expression of emotion : Acoustic Properties of Speech Are Associated With Emotional Intensity and Context. *Psychological Science, 6*(4), 219-224.

Bänziger, T., & Scherer, K.R. (2007). Using actor portrayals to systematically study multimodal emotion expression: The GEMEP corpus. *Lecture notes in computer science*, 4738, 476.

Bänziger, T., & Scherer, K.R. (2010). Introducing the Geneva multimodal emotion portrayal (GEMEP) corpus. In: K.R. Scherer, T. Bänziger, & E. B. Roesch (Eds.), *Blueprint for affective computing: A sourcebook.* Oxford, England: Oxford University Press. 271–294.

Bediou, B., Brunelin, J., D'Amato, T., Fecteau, S., Saoud M., Hénaff, M. A., et al. (2012). A comparison of facial emotion processing in neurological and psychiatric conditions. *Frontiers in Psychology, 3*, 98.

Beneke, C., Merten, J., & Krause, R. (1998) The intersubjective field: Behavioral basis of therapeutic relationships and their mental representation. *20th Ulm Workshop of Empirical Research in Psychoanalysis.*

Berntson, G.G., Boysen, S.T., & Cacioppo, J.T. (1993). Neurobehavioral Organization and the Cardinal Principle of Evaluative Bivalencea. *Annals of the New York Academy of Sciences, 702*(1), 75-102.

Bevacqua, E., Prepin, K., Niewiadomski, R., de Sevin, E., & Pelachaud, C. (2010). Greta: Towards an interactive conversational virtual companion. In: Y. Wilks & J. Benjamins (Eds.), *Artificial companions in society: Perspectives on the present and future*, 143-156.

Bevacqua, E., De Sevin, E., Hyniewska, S., & Pelachaud, C. (2012). A listener model: introducing personality traits. *Journal on Multimodal User Interfaces, 6*(1-2), 27-38.

Bianchi-Berthouze, N., & Kleinsmith, A. (2003). A categorical approach to affective gesture recognition. *Connection Science, 15*(4), 259–269.

Birdwhistell, R. (1952). *Introduction to kinesics: An annotation system for analysis of body motion and gesture.* Louisville, KY: University of Louisville.

Birdwhistell, R. (1970). *Kinesics and context*. Philadelphia: Univeristy of Pennsylvania Press.

Bonanno, G. A., Keltner, D., Noll, J.G., Putnam, F.W., Trickett, P.K., LeJeune, J., et al. (2002). When the face reveals what words do not: Facial expressions of emotion, smiling, and the willingness to disclose childhood sexual abuse. *Journal of Personality and Social Psychology, 83*(1), 94-110.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment*. New York: Basic Books.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59.

Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Attention, Perception, & Psychophysics, 65*, 7, 1078-1093.

Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R (2009) Cross-modal emotional attention: Emotional voices modulate early stages of visual processing. *Journal of Cognitive Neuroscience, 21(9)*, 1670-1679.

Brosnan, M., Ashwin, C., & Gamble, T. (2011). Greater Empathizing and reduced Systemizing in people who show a jumping to conclusions bias in the general population: Implications for psychosis. *Psychosis*, (ahead-of-print), 1-11.

Bui, T.D. (2004). Creating Emotions and Facial Expressions for Embodied Agents. PhD thesis, University of Twente, Departament of Computer Science.

Camras, L.A., Meng, Z., Ujiie, T., Dharamsi, S., Miyake, K., Oster, H., ... & Campos, J. (2002). Observing emotion in infants: facial expression, body behavior, and rater judgments of responses to an expectancy-violating event. *Emotion*, *2*(2), 179.

Carlson, J.G., & Hatfield, E. (1992). *Psychology of emotion*. Orlando: Harcourt Brace Jovanovich.

Cassell, J., Geraghty, K., Gonzalez, B., Borland, J. (2009) Modeling Culturally Authentic Style Shifting with Virtual Peers. In: *Proceedings of ICMI-MLMI*. Cambridge, MA.

Cicchetti, D.V., & Sparrow, S.A. (1981). Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American journal of mental deficiency, 86,* 127-137.

Clavel, C. Plessier, J., Martin, J.-C. Ach, L., & Morel, B. (2009). Combining facial and postural expressions of emotions in a virtual character. In: *Proceedings of Ninth International Conference on Intelligent Virtual characters*, 287-300.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement, 20,* 37-46.

Colombetti, G. & Thompson, E. (2008). The feeling body: Toward an enactive approach to emotion. *Developmental perspectives on embodiment and consciousness*, 45-68.

Coulson, M. (2004). Attributing Emotion to Static Body Postures: Recognition Accuracy, Confusions, and Viewpoint Dependence. *Journal of Nonverbal Behavior, 28*(2), 114-139.

Dael, N., Mortillaro, M., & Scherer, K.R. (2012). Emotion expression in body action and posture. *Emotion, 12*(5), 1085-1101.

Darwin, C. (1872/1998). *The Expression of the Emotions in Man and Animals*. 3rd ed. Reprint, New York: Oxford University Press.

Davidson, R., Scherer, K.R., & Goldsmith, H. (2003). *Handbook of Affective Sciences*. New York: Oxford University Press.

Davidson,R. J., Putnam,K. M., & Larson,C.L. (2000). Dysfunction in the neural circuitry of emotion regulation–a possible prelude to violence. *Science, 289*, 591–594.

De Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior*, *13*, 247–268.

Demenescu, L.R, Kortekaas, R, den Boer, J.A., Aleman, A. (2010). Impaired attribution of emotion to facial expressions in anxiety and major depression. *PLoS ONE*, *5*(12).

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

Douglas-Cowie, E., Cox, C., Martin, J.C., Devillers, L., Cowie, R., Sneddon, I., McRorie, M., Pelachaud, C., Peters, C., Lowry, O., Batliner, A. & Hönig, F. (2011). The humaine database. In: P. Petta, C. Pelachaud & R. Cowie (Eds.), *Emotion-oriented systems: The humaine handbook*. New York: Springer. 243-284.

Douglas-Cowie, E., Campbell, N., & Roach, P. (2003). Emotional speech: Towards a new generation of databases, *Speech Communication, 40*(1–2), 33–60.

Duchenne de Boulogne, G.-B. & Cuthbertson, Andrew R. (1990). *The mechanism of human facial expression*. Cambridge: Cambridge University Press.

Ekman, P. (1972). Universals and cultural differences in facial expression of emotion. In: J. R. Cole (Ed.), *Nebraska symposium on motivation*. Lincoln: University of Nebraska Press. 207–283.

Ekman, P. (1984). Expression and the nature of emotion. In: K. R. Scherer, & P. Ekman (Eds.), *Approaches to emotion*. Hillsdale, NJ: Erlbaum. 319–344.

Ekman, P. (1997). Should We Call it Expression or Communication? *Innovations in Social Science Research, 10*(4), 333-344.

Ekman, P. (1999). Basic emotions. In: T. Dalgleish, & M. Power (Eds.), *Handbook of cognition and emotion*. Chichester, England: Wiley. 45-60.

Ekman, P. (2003). Darwin, Deception, and Facial Expression. *Annals of the New York Academy of Sciences, 1000*, 205-221.

Ekman, P., & Davidson, R. J. (1994). *The nature of emotion: Fundamental questions*. New York: Oxford University Press.

Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication, 22(4),* 353-374.

Ekman, P., & Friesen, W.V. (1975). Unmasking the Face. *A guide to recognizing emotions from facial clues*. Prentice-Hall, Inc., Englewood Clis, New Jersey.

Ekman, P., & Friesen, W.V. (1982). Felt, false and miserable smiles. *Journal of Nonverbal Behavior, 6*(4), 238-252.

Ekman, P., Friesen, W.V., & Hager, J.C. (2002). *Facial Action Coding System: The Manual*. Consulting Psychologists Press.

Ellsworth, P.C. (1994). Levels of thought and levels of emotion. In: P. Ekman and R.J. Davidson (Eds), *The nature of emotion; fundamental* questions. New York: Oxford University Press. 192-196.

Ellsworth, P.C., & Smith, C.A. (1988). From appraisal to emotion: Differences among unpleasant feelings. *Motivation and Emotion, 12*, 271-302.

Ellsworth, P.C., & Tong, E.M.W. (2006). What does it mean to be angry at yourself? Categories, appraisals, and the problem of language. *Emotion, 6*, 572-586.

Fontaine, J.R., Scherer, K.R. & Roesch, E.B., & Ellsworth, P. (2007). The world of emotion is not two-dimensional. *Psychological Science*, *18*(2), 1050-1057.

Fehr, B., & Russell, J.A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General, 113*(3), 464–486.

Folkman, S., & Lazarus, R.S. (1985). If it changes it must be a process: Study of emotion and coping during three stages of a college examination. *Journal of Personality and Social Psychology, 48*, 150-170.

Frijda, N.H. (1986). *The emotions*. Cambridge, England: Cambridge University Press.

Frijda, N.H. (2000). *The psychologists' point of view. Handbook of emotions, 2*. New York: The Guilford Press. 59-74.

Frijda, N.H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, *57*, 212–228.

Frijda, N.H., Manstead, A.S.R. & Bem, S. (2000). *Emotions and beliefs: How feelings influence thoughts. Cambridge*: Cambridge University Press.

Feldman Barrett, L., & Salovey, P. (2002). *The wisdom in feeling. Psychological processes in emotional intelligence*. New York: The Guilford Press.

Gallaher, P.E. (1992). Individual differences in nonverbal behavior: Dimensions of style. *Journal of Personality and Social Psychology*, *63*(1), 133.

Gara, M.A., & Rosenberg, S. (1979). The identification of persons as supersets and subsets in free-response personality descriptions. *Journal of Personality and Social Psychology*, *37*(12), 2161.

Goffaux, V., & Rossion, B. (2006). Faces are "spatial"--holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, *32* (4), 1023.

Gomez, P., Shafy, S., & Danuser, B. (2008). Respiration, metabolic balance, and attention in affective picture processing. *Biological Psychology*, *78* (2). 138 –149.

Gosselin, P., Kirouac, G., & Doré, F.Y. (1995). Components and recognition of facial expression in the communication of emotion by actors. *Journal of Personality and Social Psychology, 68*(1), 83-96.

Grandjean, D., Sander D., Pourtois G., Schwartz S., Seghier M., Scherer K.R., et al. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature Neuroscience. 8*(2), 145-146.

Grandjean, D., Sander, D., & Scherer, K.R. (2008). Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization. *Consciousness and cognition*, *17*(2), 484.

Grizard, A., Paleari, M., & Lisetti, C.L. (2006, October). *Adaptation d'une théorie psychologique pour la génération d'expressions faciales synthétiques pour des agents d'interface*. WACA 2006, Workshop sur les Agents Conversationnels Animés, Toulouse, France.

Gross, J.J. (2002). Emotional regulation: Affective, cognitive, and social consequences. *Psychophysiology, 39*, 281-291.

Gross, J.J., & Levenson, R.W. (1997). Hiding feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Personality and Social Psychology*, *106*, 95-103.

Hager, J., & Ekman, P. (1983), The Inner and Outer Meaning of Facial Expressions. In: *Social Psychophysiology: a Sourcebook* . J.T. Cacioppo and R.E. Petty (Eds). New York: The Guilford Press.

Haidt, J., & Keltner, D. (1999). Culture and Facial Expression: Open-Ended Methods Find More Expressions and a Gradient of Recognition. *Cognition and Emotion, 13(*3), 225-266.

Harrigan, J. A. (2005). Proxemics, kinesics, and gaze. In: J.A. Harrigan, R. Rosenthal and K.R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research*. Oxford: Oxford University Press.

Hess, U., Adams, R.B.J., & Kleck, R.E. (2004). *Facial Appearance, Gender, and Emotion Expression. Emotion, 4, 378-388.*

Hess, U., Banse, R., & Kappas, A. (1995). The intensity of facial expression is determined by underlying affective state and social situation. *Journal of Personality and Social Psychology, 69*, 280-288.

Hyniewska, S., Niewiadomski, R., Mancini, M., & Pelachaud, C. (2010). *Expression of affects in Embodied Conversational Agent.* In: Scherer, K.R., Banziger, T., Roesch, E. (Eds.), A Blueprint for an affectively competent agent: Cross-fertilization between Emotion Psychology, Affective Neuroscience, and Affective Computing. Oxford: Oxford University Press. *213-221.*

Isaacowitz, D.M., Löckenhoff, C.E., Lane, R. D., Wright, R., Sechrest, L., Riedel, R., & Costa, P. T. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging, 22,* 147-159.

Itier, R.J., Alain, C. Sedore, K., & McIntosh, A.R. (2007). Early face specificity: it's in the eyes! *. Journal of Cognitive Neuroscience, 19*(11), 1815-1826.

Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin, 115,* 288–299.

James, W. (1884). What is an emotion? *Mind, 9,* 188–205.

James, W. (1890/1981). *The principles of psychology*. Cambridge, MA: Harvard University Press.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics, 14*, 201-211.

Johnstone, T., van Reekum, Carien, M., Hird, K., Kirsner, K., & Scherer, K.R. (2005). Affective speech elicited with a computer game. *Emotion*, 5(4), 513-518.

Juslin, P.N., & Scherer, K.R. (2005). Vocal expression of affect. In: J. Harrigan, R. Rosenthal & K.R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research.* Oxford: Oxford University Press.

Kang, S.H., & Gratch, J. (2010). *Virtual humans elicit socially vocal expression of affect verbal self disclosure*. 23rd international conference on computer animation and social agents, Saint-Malo, France.

Kanwisher, N., McDermott, J., & Chun, M.M (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience, 17(11),* 4302.

Kafetsios, K., & LaRock, E. (2005). Cognition and Emotion: Aristotelian affinities with contemporary emotion research. *Theory & Psychology, 15*, 5, 639-657.

Kaiser, S., & Scherer, K.R. (1998). Models of 'normal' emotions applied to facial and vocal expressions in clinical disorders. In: W. F. Flack, Jr. & J. D. Laird (Eds.), *Emotions in Psychopathology*. New York: Oxford University Press. 81-98.

Kaiser, S., & Wehrle, T. (2001). Facial expressions as indicators of appraisal processes. In: K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotions: Theory, methods, research*. New York: Oxford University Press. 285-300.

Kaiser, S., & Wehrle, T. (2006). Modeling appraisal theory of emotion and facial expression. In: N. Magnenat-Thalmann, D. Pai, A. Paiva, & E. Wu (Eds.), In: *Proceedings of the 19th International Conference on Computer Animation and Social Agents* (CASA 2006). Computer Graphics Society (CGS). 121-130.

Kaiser, S., Wehrle, T., & Schmidt, S. (1998). Emotional episodes, facial expression and reported feelings in human-computer interactions. In: N.H. Frijda (Ed.), *Proceedings of the Xth Conference of the International Society for Research on Emotions*, ISRE Publications.

Kappas, A., & Pecchinenda, A. (1999). Don't wait for the monsters to get you: A video game task to manipulate appraisals in real time. *Cognition and Emotion, 13*, 119-124.

Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame, *Journal of Personality and Social Psychology, 68*, 441–454.Kleinsmith, A., & Bianchi-Berthouze, N., (2007). Recognizing Affective Dimensions from Body Posture. In : *Proceedings of Second International Conference on Affective Computing and Intelligent Interaction*, 48-58.

Kleinsmith, A., De Silva, R., Bianchi-Berthouze, N. (2006). Cross-Cultural Differences in Recognizing Affect from Body Posture. *Interacting with Computers, 18* (6), 1371-1389.

Kohler, C.G., Turner, T. H., Gur, R.E., & Gur, R.C. (2004). Recognition of facial emotions in neuropsychiatric disorders. *CNS Spectrums*. *9*, 267–274.

Kraut, R.E. (1982). Social pressure, facial feedback, and emotion. *Journal of Personality and Social Psychology*, *42*, 853-863.

Kraut, R.E., & Johnston, R.E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology, 37*, 1539-1553.

Kuppens, P. (2010). From appraisal to emotion. *Emotion Review, 2*, 157-158.

Kuppens, P., Van Mechelen, I., Smits, D.J.M. & De Boeck, P. (2003). The appraisal basis of anger: Specificity, necessity and sufficiency of components. *Emotion*, *3*(3), 254.

Laban, R. (1956). *Principles of Dance Movement Notation: With 114 Basic Movement Graphs and Their Explanation*. London: Macdonald & Evans.

Laban, R. (1975). Principles of dance and movement notation (Vol. 2). London: MacDonald & Evans.

Laban, R. (1928). Schrifttanz. *Wien: Universal Edition*.

Labukt, I. (2012). Hedonic Tone and the Heterogeneity of Pleasure. *Utilitas*, *24*, 172-199.

Laird, J.D. (1974). Self-attribution of emotion: The effects of expressive behavior on the quality of emotional experience. Journal of Personality and Social Psychology, 29, 475-486.

Lance, B. & Marsella, S. (2007). Emotionally Expressive Head and Body Movements during Gaze Shifts. In: *Proceedings of Seventh International Conference on Intelligent Virtual characters*, 72-85.

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

Laukka, P. (2005). Categorical perception of vocal emotion expressions. *Emotion*, *5*(3), 277-95.

Lazarus, R.S. (1984). On the primacy of cognition. *American Psychologist*, *39*(2), 124-129.

Lazarus, R.S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

MacDowell, K.A., & Mandler, G. (1989). Constructions of emotion: Discrepancy, arousal, and mood. *Motivation and Emotion, 13,* 105-125.

Malatesta, L., Raouzaiou, A., Karpouzis, K., & Kollias, S.D. (2009). Towards Modeling Embodied Conversational Agent Character Profiles Using Appraisal Theory Predictions in Expression Synthesis. *Applied Intelligence, 30*(1), 58-64.

Malatesta, L., Caridakis, G., Raouzaiou, A., & Karpouzis, K. (2007). Agent Personality Traits in Virtual Environments Based on Appraisal Theory Predictions. In: *Artificial and Ambient Intelligence, Language, Speech and Gesture for Expressive Characters*, Newcastle upon Tyne, UK.

Manstead, A.S.R., & Fischer, A.H. (2001). Social appraisal: The social world as object of and influence on appraisal processes. In: K.R. Scherer, A. Schorr, & T. Johnston (Eds.), *Appraisal processes in emotion: Theory, methods, research*. New York, NY: Oxford University Press.

Niewiadomski, R., Hyniewska, S., & Pelachaud, C. (2009). Evaluation of Multimodal Sequential Expressions of Emotions in ECA. In: *Proceedings of the Third International Conference on Affective Computing and Intelligent Interaction*, 635-641.

Niewiadomski, R., Hyniewska, S., & Pelachaud, C. (2011). Constraint-Based Model for Synthesis of Multimodal Sequential Expressions of Emotions. *IEEE Transactions in Affective computing, 2*(3), 134-146.

Niewiadomski, R., & Pelachaud, C. (2010). Affect expression in ECAs: Application to politeness displays. *International Journal of Human-Computer Studies*, *68*(11), 851-871.

Öhman, A. (2005). Conditioned fear of a face: A prelude to ethnic enmity? *Science, 309*, 711-712.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General, 130*, 466-478.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. New York: Cambridge University Press.

Ortony, A., & Turner, T. (1990). What's basic about basic emotions? *Psychological Review, 97,* 315-331.

Paleari, M., Grizard, A., & Lisetti, C.L. (2006). Adapting Psychologically Grounded Facial Emotional Expressions to Different Anthropomorphic Embodiment Platform. In: *Proceedings of Florida Artificial Intelligence Research Society*. Key West, Florida.

Pan, X., Gillies, M., Sezgin, T.M., & Loscos, C. (2007).Expressing Complex Mental States through Facial Expressions. In: *Proceedings of Second International Conference on Affective Computing and Intelligent Interaction*, 745-746.

Parkinson, B. (1997). Untangling the appraisal–emotion connection. *Personality and Social Psychology Review, 1*, 62–79.

Parkinson, B., & Manstead, A.S.R. (1993). Making sense of emotion in stories and social life. *Cognition and Emotion, 7*, 295-323.

Pollick, F. E. (2004). The features people use to recognize human movement style. In: *Gesture-Based Communication in Human-Computer Interaction*, 10-19. Springer Berlin Heidelberg.

Pollick, F.E., Paterson, H., Bruderlin, A., & Sanford, A.J. (2001) Perceiving affect from arm movement. *Cognition, 82*(2), 51-61.

Plutchik, R. (1962). *The emotions: Facts, theories, and a new model*. New York: Random House.

Plutchik, R. (1980), Emotion: Theory, research, and experience: Vol. 1. *Theories of emotion, 1.* New York: Academic.

Reisenzein, R. (1994) Pleasure-arousal Theory and the Intensity of Emotions *Journal of Personality and Social Psychology, 67*, 525–39.

Reisenzein, R. (2000). The subjective experience of surprise. In: H. Bless & J. P. Forgas (Eds.), The message within: The role of subjective experience in social cognition and behavior . Philadelphia: Psychology Press. 262–279.

Reisenzein, R., & Schoenpflug, W. (1992). Stumpf's cognitiveevaluative theory of emotion. *American Psychologist, 47*, 34–45.

Roesch, E.B., Tamarit, L., Reveret, L., Grandjean, D., Sander, D., & Scherer, K.R. (2010). FACSGen: A tool to generate synthetic, realistic, static and dynamic, FACS-based facial expressions. *Journal of Nonverbal Behavior*.

Roether, C.L., Omlor, L., Christensen, A., & Giese, M.A. (2009). Critical features for the perception of emotion from gait. *Journal of Vision*, 9(6), 1–32.

Rosch, E. (1978). Principles of categorization. In: E. Rosch (Ed.), *Cognition and categorization*. Hillsdale, N.J.: Erlbaum. 28-49.

Roseman, I.J. (1984). Cognitive determinants of emotion: A structural theory. In: Shaver, P. (Ed.), *Review of personality and social psychology, 5*. Beverly Hills, CA: Sage. 11-36.

Roseman, I.J. (1991). Appraisal determinants of discrete emotions. *Cognition and Emotion, 5*, 161-200.

Roseman, I.J., Spindel, M.S., & Jose, P.E. (1990). Appraisals of emotion-eliciting events: Testing a theory of discrete emotions. *Journal of Personality and Social Psychology, 59*, 899–915.

Roseman, I.J., & Smith, C.A. (2001). Appraisal theory. In: K.R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research*. Oxford: Oxford University Press. 3-19.

Russell, J.A., Bachorowski, J-A., & Fernández-Dols, J.M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54*, 329–349.

Russell, J.A., & Fehr, B. (1994). Fuzzy concepts in a fuzzy hierarchy: Varieties of anger. *Journal of Personality and Social Psychology, 67*, 186–205.

Ruttkay, Z. (2001). Constraint-Based Facial Animation. In: *International Journal of Constraints, 6*, 85-113.

Ruttkay, Z., Noot, H., & ten Hagen, P. (2003). Emotion Disc and Emotion Squares: tools to explore the facial expression face, *Computer Graphics Forum*, *22* (1), 49-53.

Salovey, P., & Mayer, D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality, 9*, 185-211.

Sander, D., Grafman, J., & Zalla, T. (2003). The human amygdala: An evolved system for relevance detection. *Reviews in the Neurosciences, 14*(4), 303 -316.

Sander, D., Grandjean, D., Kaiser, S., Wehrle, T., & Scherer, K.R. (2007). Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *European Journal of Cognitive Psychology*, *19*(3), 470-480.

Sander, D., Grandjean, D., & Scherer, K.R. (2005). A systems approach to appraisal mechanisms in emotion. *Neural Networks. 18*, 317-352.

Sato, W., Kochiyama, T. & Yoshikawa, S. (2011). The inversion effect for neutral and emotional facial expressions on amygdala activity. *Brain Research, 1378,* 84-90.

Sato, W. & Yoshikawa, S. (2007). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition, 104*(1), 1-18.

Schachter, S., & Singer, J.E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69*, 379-399.

Scheflen, A.E. (1964). The significance of posture in communication systems. *Psychiatry, 27,* 316–331.

Scherer, K.R. (1984). On the nature and function of emotion: A component process approach. In: Scherer, K.R., & Ekman, P. (Eds.), *Approaches to emotion*. Hillsdale, NJ: Erlbaum. 293-317.

Scherer, K.R. (1985). Vocal affect signaling: A comparative approach. In: J. Rosenblatt, C. Beer, M.-C. Busnel, & P.J.B. Slater (Eds.), *Advances in the study of behaviour, 15*. New York: Academic Press. 189-244.

Scherer, K.R. (1992). What does facial expression express? In: K. Strongman (Ed.), *International review of studies on emotion, 2*. Chichester, England: Wiley. 139-165.

Scherer, K.R. (1993). Studying the emotion-antecedent appraisal process: An expert system approach. *Cognition and Emotion, 7*, 325-255.

Scherer, K.R. (2000). Emotion. In: M. Hewstone, Stroebe, W. (Eds.) *Introduction to Social Psychology: A European perspective* (3rd. ed), 151-191.

Scherer, K.R. (2001). Appraisal considered as a process of multi-level sequential checking. In: K. R. Scherer, A. Schorr, & T. Johnstone (Eds.). *Appraisal processes in emotion: Theory, Methods, Research*. New York and Oxford: Oxford University Press. 92-120.

Scherer, K.R. (2004). Ways to Study the Nature and Frequency of Our Daily Emotions: Reply to the Commentaries on “Emotions in Everyday Life”. *Social Science Information, 43*(4), 667-689.

Scherer, K.R. (2005). What are emotions? And how can they be measured? *Social Science, Information, 44*. 695-729.

Scherer, K.R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion, 23*(7), 1307-1351.

Scherer, K.R., & Ceschi, G. (1997). Lost luggage: A field study of emotion-antecedent appraisal. *Motivation and Emotion, 21*, 211-235.

Scherer, K.R., & Ceschi, G. (2000). Criteria for emotion recognition from verbal and nonverbal expression: Studying baggage loss in the airport. *Personality and Social Psychology Bulletin. 26*(3), 327-339.

Scherer, K.R, & Ellgring, H. (2007). Are Facial Expressions of Emotion Produced by Categorical Affect Programs or Dynamically Driven by Appraisal? *Emotion, 7*, 113-130.

Scherer, K.R., Zentner, M.R., & Stern, D. (2004). Beyond surprise: the puzzle of infants' expressive reactions to expectancy violation. *Emotion*, *4*(4), 389.

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., & Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, *44*(5), 718-730.

Schorr, A. (2001). Appraisal—the evolution of an idea. In: K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research*. New York: Oxford University Press. 20–34.

Schröder, M., Pammi, S., & Türk, O. (2009). Multilingual MARY TTS participation in the Blizzard Challenge 2009. In: *Blizzard Challenge 2009*. Edinburgh, UK.

Shafer, A.T., Matveychuk, D., Penney, T., O'Hare, A.J., Stokes, J., & Dolcos, F. (2012). Processing of Emotional Distraction is both Automatic and Modulated by Attention: Evidence from an Event-Related fMRI Investigation. *Journal of Cognitive Neuroscience, 24(*5), 1233–1252.

Shiota, M.N., Campos, B., & Keltner, D. (2003). The faces of positive emotion: Prototype displays of awe, amusement, and pride. *Annals of the New York Academy of Sciences, 1000,* 296-299.

Siemer, M., Mauss, I., & Gross, J.J. (2007). Same situation–different emotions: How appraisals shape our emotions. *Emotion*, *7*(3), 592-600.

Smith, C.A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology, 48*, 813-838.

Smith, C.A., & Pope, L.K. (1992). Appraisal and emotion: The interactional contributions of dispositional and situational factors. In: M.S. Clark (Ed.), *Review of personality and social psychology*, *14*: Emotion and social behavior. Newbury Park, CA: Sage. 32–62.

Smith, C. A., Scott, H.S. (1997). A componential approach to the meaning of facial expression. In: J.A., Russell, & J., Fernández-Dols (Eds.), *The psychology of facial expression*. Cambridge: Cambridge University Press. 229-254.

Soussignan, R. (2002). Duchenne smile, emotional experience and autonomic reactivity: A test of the facial feedback hypothesis. *Emotion*, *2*, 52-74.

Soussignan, R., Chadwick, M., Philip, L., Conty, L., Dezecache, G., & Grèzes, J. (2012). Self-Relevance Appraisal of Gaze Direction and Dynamic Facial Expressions: Effects on Facial Electromyographic and Autonomic Reactions.

De Spinoza, B. (1955). The Chief Works (R.H.M. Elwes, Trans.) New York: Dover. (Original work published 1677).

Stoiber, N., Seguier, R., & Breton, G. (2009). Automatic Design of a Control Interface for a Synthetic Face. In: *Proceedings of International Conference on Intelligent User Interfaces*, 207-216.

Strack, F., Martin, L.L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobstrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *54*, 768-777.

Taylor, M. J., George, N., & Ducorps, A. (2001). Magnetoencephalographic evidence of early processing of direction of gaze in humans. *Neuroscience Letters*, *316*, 173–177.

Tomkins, S.S. (1962). *Affect, imagery, consciousness: The positive affects (Vol. 1)*. New York: Springer.

Tomkins, S. S. (1982). The face of man: Expressions of universal emotions in a New Guinea village: by Paul Ekman. Garland STPM, NY, 1980. *Social Science & Medicine*, *16*(7), 844-845.

Tomkins, S. S. (1984). Affect theory. *Approaches to emotion*, *163*, 195.

Tomkins, S.S., & McCarter, R. (1964). What and where are the primary affects? Some evidence for a theory. *Perceptual and Motor Skills, 18I,* 119-158.

Tourangeau, R., & Ellsworth, P.C. (1979). The role of facial response in the experience of emotion. *Journal of Personality and Social Psychology*, *37*, 1519–1531.

Trout, D.L., & Rosenfeld, H.M. (1980). The effect of postural lean and body congruence on the judgment of psychotherapeutic rapport. *Journal of Nonverbal Behavior*, *4*(3), 176-190.

Tsapatsoulis, N., Raouzaiou, A., Kollias, S., Crowie, R., & Douglas-Cowie, E. (2002). Emotion Recognition and Synthesis Based on MPEG-4 FAPs. In: I. Pandzic, R. Forchheimer (Eds.), *MPEG-4 Facial Animation - The standard, implementations, applications*. UK: John Wiley & Sons.

Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, *7*, 487–494.

Vrtička, P., & Vuilleumier, P. (2012). Neuroscience of human social interactions and adult attachment style. *Frontiers in Human Neuroscience*, *6*.

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*, 585–594.

Wagner, H.L., & Smith, J. (1991). Facial expression in the presence of friends and strangers. *Journal of Nonverbal Behavior*, *15*(4), 201-214.

Wilhelm, P. & Schoebi, D. (2007). Assessing Mood in Daily Life: Structural Validity, Sensitivity to Change, and Reliability of a Short -Scale to Measure Three Basic Dimensions of Mood. *European Journal of Psychological Assessment*, *23*, 258-267.

Wallbott, H.G. (1985). Hand movement quality: A neglected aspect of nonverbal behavior in clinical judgment and person perception. *Journal of Clinical Psychology*, *41(3)*, 345-359.

Wallbott, H.G. (1998). Bodily expression of emotion. *European Journal of Social Psychology, 28,* 879-896.

Wallbott, H.G., & Scherer, K.R. (1986). Cues and channels of emotion recognition. *Journal of Personality and Social Psychology, 51(4),* 690-699.

Whalen, P.J., Rauch, L.S., Etcoff, N.L., Mclnerney, S.C., Lee, M.B., & Jenike, M.A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *The Journal of Neuroscience*, *18*(1), 411–418.

Whalen, P.J., Kagan, J., Cook, R.G., Davis, F.C., Kim, H., Polis, S., McLaren, D. G., Somerville, L. H., McLean, A. A., Maxwell, J. S., & Johnstone, T. (2004). Human amygdala responsivity to masked fearful eye whites. *Science*, *306*(5704), 2061.

Whissell, C. (1989). The Dictionary of Affect in Language, In: Plutchik, R., Kellerman, H., (Eds.), *Emotion: Theory, Research, and Experience, volume 4: The Measurement of Emotions*. San Diego, USA: Academic Press, Inc. 113–131.

With, S., & Kaiser, S. (2011). Sequential Patterning of Facial Actions in the Production and Perception of Emotional Expressions. *Swiss Journal of Psychology, 70*(4), 241-252.

Wundt, W. (1874/1902/1904). *Principles of physiological psychology* (Edward Bradford Titchener, Trans.) from the 5th German ed., published 1902; 1st German ed. published 1874.

Yantis, S. (1998). Control of visual attention. In: H. Pashler (Ed.), *Attention*. Hove, UK: Psychology Press. 223-256.

Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum.

Zajonc, R.B. (1980). Thinking and feeling: Preferences need no inferences. *American Psychologist*, *35*, 151–175.

Appendix I

Guidelines presented to participants in the *judgement of the one minute extracts from the Lost Luggage corpus*

Vous allez voir plusieurs vidéos, chacune d'une minute, et les annoter sur un programme (Anvil).

Votre tâche est d'indiquer les changements entre états mentaux de la personne. Sélectionnez une période (en cliquant **start** sur une image se situant vers le début de l'état, puis **end** vers sa fin) et essayer de définir cet état par un mot :

Ex. rage, Questionnement, Anti-social

ou un groupe de mots :

Ex :
La personne émet un message de détresse pour qu'on la réassure
La personne veut savoir ce qui se passe
La personne est furieuse
La personne trouve ce qu'elle voit déplaisant

Un film peut contenir plusieurs états émotionnels et des moments de neutralité. Indiquez les.

Grand merci !

Appendix II

Set of evaluative questions for each video clip presented to participants from the *Naturalistic expression study*

Dans l'extrait que vous venez de voir, avez-vous l'impression que la personne :	FORTEMENT EN DESACCORD	EN DESACCORD	LEGEREMENT EN DESACCORD	NI EN DESACCORD NI EN ACCORD	LEGEREMENT EN ACCORD	EN ACCORD	FORTEMENT EN ACCORD
vient d'être confrontée à un élément soudain ?	○	○	○	○	○	○	○
vient d'être confrontée à un événement qui pourrait l'empêcher de réaliser ses plans ?	○	○	○	○	○	○	○
vient d'être confrontée à un événement important et inattendu ?	○	○	○	○	○	○	○
se sent capable de maîtriser la situation dans laquelle elle se trouve ?	○	○	○	○	○	○	○
considère la situation, dans laquelle elle se trouve à cause d'autrui, comme inacceptable ?	○	○	○	○	○	○	○
croit se comporter de façon appropriée à l'événement, en respectant ses rêgles morales ?	○	○	○	○	○	○	○
semble impliquée dans la discussion ?	○	○	○	○	○	○	○
veut établir ou garder une bonne relation avec l'interlocuteur ?	○	○	○	○	○	○	○
exprime franchement ce qu'elle ressent?	○	○	○	○	○	○	○

Figure 5. First page presented to participants after each video: sentences evaluating appraisal outcomes using likert scales

*Evaluez l'intensité des émotions ressenties par la personne sur les échelles allant de 0 (pas de ressenti de l'émotion donnée) à 6 (intensité très forte)

	0 absence de l'émotion	1	2	3	4	5	6 intensité forte
Joie	○	○	○	○	○	○	○
Colère	○	○	○	○	○	○	○
Soulagement	○	○	○	○	○	○	○
Tristesse	○	○	○	○	○	○	○
Mépris	○	○	○	○	○	○	○
Peur	○	○	○	○	○	○	○
Honte	○	○	○	○	○	○	○

Figure 6. Second page presented to participants after each video: emotion labels using likert scales

Appendix III
Naturalistic expression study scores for each video

Standard-deviations for *Naturalistic expression study* scores for each video

Video	Obs.	q1	q2	q3	q4	q5	q6	q91	q92	q93	q94	q95	q96	q97
77	23	1,49	1,33	1,66	1,79	1,41	1,06	0,50	0,80	1,44	1,43	0,86	1,47	1,08
04a	16	1,59	1,59	1,69	1,78	1,57	1,61	0,34	1,00	0,89	1,63	1,82	1,98	0,96
103a	17	1,64	1,62	1,50	1,17	1,49	1,10	0,59	0,47	1,21	0,47	1,30	1,29	0,72
103b	14	0,27	1,70	0,47	2,11	2,02	1,99	0,94	1,91	1,02	0,74	1,95	1,09	0,58
108a	17	1,80	1,95	1,52	1,37	1,62	1,14	0,86	0,72	1,17	0,81	0,77	0,77	0,72
108b	19	1,71	1,58	1,52	1,22	1,52	1,21	0,90	0,70	1,34	1,00	0,82	0,87	0,67
10a	14	1,65	1,94	1,82	1,70	1,46	1,52	0,58	2,18	0,94	1,48	1,75	1,79	1,98
10b	77	1,87	1,78	1,70	1,77	1,70	1,32	0,19	1,73	0,67	1,78	1,79	1,82	1,16
10c	40	1,61	1,54	1,51	1,55	1,42	1,17	0,16	1,32	0,40	1,84	1,42	2,00	1,28
110a	19	1,73	1,52	1,68	1,70	1,72	1,86	0,50	2,00	0,23	1,71	2,16	1,89	0,93
110b	19	2,27	2,36	2,38	1,93	2,27	1,80	0,00	2,38	0,00	1,49	2,13	0,96	1,12
110c	17	2,18	2,21	2,12	1,90	1,79	2,03	0,00	2,18	0,00	1,00	1,60	0,87	1,12
111a	39	1,64	1,56	1,57	1,47	1,58	1,43	0,16	1,34	0,43	1,37	1,79	1,60	1,10
17a	20	1,49	1,84	1,98	1,81	2,01	1,57	0,31	1,41	0,94	1,47	1,26	1,50	0,73
17b	19	1,29	1,43	1,31	1,71	1,64	1,35	0,00	1,47	0,32	1,94	0,76	1,61	0,65
19a	19	1,41	1,55	1,41	1,61	1,40	1,17	0,90	1,54	0,23	0,50	1,80	1,05	0,96
19b	11	1,49	1,85	1,48	1,81	2,15	1,97	0,47	0,82	0,93	0,50	0,65	1,21	0,65
19c	19	1,65	1,37	1,71	1,39	1,28	1,81	0,32	1,66	0,84	0,93	1,55	1,15	1,54
33a	14	1,34	1,03	1,54	2,09	1,74	1,53	0,36	1,25	0,63	1,69	1,41	1,95	1,03
45a	73	1,76	1,73	1,75	1,53	1,49	1,40	0,67	1,83	0,46	1,86	1,95	2,04	1,33
50a	17	1,66	1,58	1,59	1,62	2,01	1,48	0,56	1,05	1,45	0,72	1,91	1,03	1,34
51a	12	2,04	2,12	2,01	1,21	1,66	1,93	0,00	1,28	0,29	2,14	1,95	2,39	1,57
51b	41	1,92	1,26	2,36	1,42	1,35	1,53	2,11	1,28	2,00	1,45	1,41	1,86	1,75
60a	40	1,59	1,66	1,38	1,32	1,75	1,46	0,16	1,30	0,76	1,33	1,93	1,20	1,58
60b	77	1,61	1,51	2,05	1,55	1,80	1,82	1,68	1,73	1,14	1,63	1,86	1,68	1,34
60c	20	1,12	1,62	1,39	1,59	1,75	1,79	0,89	0,89	0,49	1,77	1,07	1,82	2,15
60s	14	1,41	1,28	1,19	1,44	2,05	1,22	0,27	1,70	0,47	2,11	2,02	1,99	1,95
62a	55	2,53	2,56	2,04	1,73	2,28	1,71	0,73	1,55	0,59	1,93	1,77	1,98	1,95
62b	77	2,14	1,61	1,67	1,56	1,65	1,50	0,16	1,25	1,91	1,74	1,92	1,63	1,32
66a	52	2,36	2,38	1,69	2,25	1,81	1,53	1,99	1,89	1,29	1,47	1,81	1,69	1,75
66b	40	1,57	1,84	1,75	1,53	1,83	1,38	0,00	1,83	0,65	1,44	1,93	1,53	1,69
66b_c	14	1,45	1,56	1,53	1,67	1,91	1,76	1,67	1,57	1,95	0,95	1,83	0,89	0,84
68a	19	1,74	1,45	1,65	1,50	1,61	1,65	0,73	0,95	0,96	0,68	1,06	1,03	0,96
68b	19	1,59	1,61	1,67	1,71	1,64	1,57	0,76	1,57	1,27	0,70	1,56	1,20	1,01
68c	17	1,58	1,30	1,27	1,33	1,33	1,41	1,74	0,66	1,67	0,44	0,75	0,33	1,12
70a	77	2,04	1,68	2,07	1,78	1,59	1,40	1,86	1,22	1,83	0,94	1,52	1,40	1,11
70b	17	1,62	1,31	1,62	1,62	1,25	1,82	0,80	1,82	0,94	1,66	1,20	1,46	1,05
76a	40	1,60	1,77	1,45	1,69	1,50	1,30	0,45	0,81	1,22	1,47	1,08	1,48	1,41
76c	19	1,57	1,54	1,25	1,56	1,08	1,44	0,84	0,37	1,50	1,02	0,54	0,77	0,83

Note. The first line of the Table states labels of likert scales used by participants. Their codes are as follows: q1= suddenness, q2= obstruction, q3= discrepancy, q4= coping potential, q5= respect ofinternal standards, q6= external standard violation, q91=Joy, q92=Anger, q93= Relief, q94= Sadness, q95= Contempt, q96= Fear, q97= Shame.

Means for Naturalistic expression study scores for each video

Video	Obs.	q1	q2	q3	q4	q5	q6	q91	q92	q93	q94	q95	q96	q97
77	23	4,27	3,95	3,91	3,95	3,00	5,45	0,18	0,41	1,18	1,36	0,45	1,50	0,73
04a	16	5,38	5,38	4,94	4,38	3,25	4,94	0,13	1,06	0,63	2,00	1,69	2,06	0,63
103a	17	4,24	3,35	3,41	5,65	3,29	5,71	0,29	0,29	1,71	0,29	0,76	0,82	0,47
103b	14	0,07	1,57	0,29	2,00	1,36	2,57	0,50	1,57	0,57	0,36	2,57	0,50	0,21
108a	17	3,65	3,94	4,06	4,59	3,65	5,06	0,35	0,47	0,88	0,82	0,29	0,71	0,53
108b	19	4,47	3,95	3,89	5,47	3,26	5,63	0,47	0,47	0,84	0,68	0,68	0,74	0,32
10a	14	3,43	3,93	3,64	4,57	3,14	5,00	0,21	1,86	0,57	1,21	1,86	1,50	1,07
10b	77	4,53	4,51	4,25	3,70	3,99	4,69	0,04	1,57	0,31	1,66	1,43	2,06	0,92
10c	40	4,90	4,95	4,98	3,15	3,65	5,23	0,03	1,30	0,13	2,18	1,03	2,08	1,00
110a	19	4,74	5,26	4,53	3,32	4,95	4,32	0,16	2,74	0,05	1,37	2,74	1,42	0,74
110b	19	3,95	3,84	3,74	3,21	3,79	3,37	0,00	2,32	0,00	0,89	1,89	0,53	0,84
110c	17	2,88	3,00	3,12	2,65	2,71	2,88	0,00	1,65	0,00	0,65	0,94	0,41	1,00
111a	39	4,44	4,69	4,41	3,49	3,97	4,44	0,03	1,69	0,15	1,26	1,79	1,54	0,54
17a	20	5,00	4,70	4,35	4,00	3,05	5,40	0,10	1,00	0,60	1,45	0,70	2,15	0,30
17b	19	5,68	5,58	5,53	2,63	3,63	4,42	0,00	0,95	0,11	2,26	0,37	3,37	0,26
19a	19	5,26	5,21	5,11	4,63	5,79	5,47	0,53	2,63	0,05	0,37	2,37	0,89	0,63
19b	11	5,27	4,73	5,00	3,91	3,73	4,45	0,27	0,55	0,55	0,64	0,27	1,55	0,27
19c	19	3,79	4,26	3,84	5,05	4,74	4,95	0,11	1,89	0,42	0,74	1,21	1,11	1,47
33a	14	2,64	6,14	2,93	4,29	4,64	4,21	0,14	0,79	0,36	1,93	0,86	2,50	0,86
45a	73	4,85	4,95	4,71	2,60	3,97	4,11	0,14	1,79	0,11	1,99	1,78	2,86	0,89
50a	17	4,65	4,12	4,18	5,00	4,06	4,94	0,24	1,12	0,88	0,53	1,47	1,06	1,06
51a	12	5,00	4,83	4,75	2,00	3,25	3,92	0,00	1,00	0,08	1,75	1,17	2,58	1,50
51b	41	4,15	4,83	3,46	3,02	4,02	4,56	1,29	1,00	1,37	0,90	1,10	2,17	1,68
60a	40	3,65	3,83	3,70	5,00	3,58	4,60	0,03	1,50	0,30	0,98	1,90	0,83	1,03
60b	77	4,60	5,30	4,12	2,84	4,48	3,60	0,68	2,40	0,51	2,05	2,21	1,92	0,77
60c	20	5,75	5,10	4,95	3,10	3,30	4,05	0,20	0,80	0,15	1,75	0,75	2,55	1,90
60s	14	5,14	5,36	5,21	3,07	3,29	4,64	0,07	1,57	0,29	2,00	1,36	2,57	1,36
62a	55	3,95	4,13	4,49	2,27	3,40	4,16	0,20	2,31	0,15	3,05	1,60	2,22	1,69
62b	77	4,82	5,16	4,90	2,64	4,30	4,29	0,03	1,75	0,92	3,32	2,08	2,10	0,86
66a	52	3,52	3,75	4,56	3,56	4,06	4,83	1,63	2,00	0,88	1,15	1,21	1,67	1,90
66b	40	4,72	4,90	4,68	3,38	4,30	4,43	0,00	2,03	0,13	1,23	1,60	1,40	1,40
66b_c	14	4,64	4,14	3,79	4,79	3,57	5,21	1,21	1,00	1,64	0,86	0,86	0,79	0,64
68a	19	4,37	3,68	4,21	4,16	3,58	3,95	0,26	0,68	0,63	0,37	0,68	0,95	0,42
68b	19	4,74	4,42	4,37	3,84	3,63	4,16	0,63	1,32	0,95	0,47	1,00	0,89	0,63
68c	17	3,41	3,06	3,00	5,41	2,47	5,00	2,82	0,24	1,82	0,24	0,24	0,12	0,59
70a	77	3,61	4,21	3,49	5,05	3,74	5,25	0,99	1,44	1,18	0,64	1,29	1,12	0,65
70b	17	3,53	4,71	5,00	4,88	4,06	0,94	0,41	1,94	0,47	1,59	0,94	1,59	0,88
76a	40	4,48	4,20	4,55	4,68	3,10	5,48	0,18	0,63	0,73	1,43	0,65	1,28	0,90
76c	19	4,58	3,84	4,32	4,74	2,95	5,21	0,42	0,16	1,37	0,58	0,21	0,53	0,37

Note. The first line of the Table states labels of likert scales used by participants. Their codes are as follows: q1= suddenness, q2= obstruction, q3= discrepancy, q4= coping potential, q5= respect ofinternal standards, q6= external standard violation, q91=Joy, q92=Anger, q93= Relief, q94= Sadness, q95= Contempt, q96= Fear, q97= Shame.

Appendix IV
Body action codes

A scheme was created to code for body actions and postures that are not annotated in the FACS (Ekman, Friesen & Hager, 2002). The aim of the scheme is to pick up changes visible by the eye that could be relevant for non guided exploration of emotional displays. It is suited particularly for the audio-visual corpus we extracted based on the Lost Luggage case (Scherer & Ceschi, 1997). It presents a face to face interaction between a claiming person (a passenger) and an interlocutor (hostess). Only the behaviour of the passenger is annotated.
The changes in the positioning are noted for the passenger relative to the hostess.
Given that the passenger is in a sitting position in front of a desk, and in the majority of cases facing the hostess, front and backward movements are not relative to the vertical axis of the passenger, but in relation to the hostess. Thus, a backward movement is a movement away from the hostess and forward is towards the hostess.

We annotated eye gaze, torso, arms, hands and shoulders movements, body action-positions such as keeping elbows on the desk, closing the hand into a fist, manipulators (hand level, face level, other), breathing, shaking and fidgeting, illustrators, beats.

Eye Gaze:
To complete the FACS codes for the eyes (left, right, up, down gazes), we created several additional codes. These were looking in the eyes of the hostess (or mutual gaze; 301), looking at the hostess in a less specific way (300), looking straight (302).

Torso:
Torso changes have to be annotated when there is a shift away from the neutral position in the torso. It could be a change perceived as initiated during the film (a clip starting with a neutral position than a movement away from the neutrality), one maintained during the entire clip (posture kept shifted away from neutrality during the whole duration of the clip) or an ending one (movement of a comeback to neutrality). All the three are coded in the same way, undifferentiated, with the same codes.

The actions/postures that could be observed as happening are **leans of the trunk,** and a **slouched (upper body collapsed)** or an **erect position.**

A lean refers to the angle of the trunk with respect to a vertical line drawn from the midline of the head and chest, to the hips.
No lean is recorded when the trunk is upright, id est head and shoulder in a direct vertical line over the hips.

Neural torso position

A forward lean (400) implies an inclination to the front, i.e. head and shoulders backward of upright relative to the hips (Harrigan, 2005). In other words shoulders are thrust closer to the hostess and lowered closer to the desk, unless compensated by another movement.

Forward lean

A *backward lean* (401) implies a retreat, i.e. head and shoulders backward of upright relative to the hips. In other words it is a movement away from the hostess, with shoulders that normally follow, unless a compensation by a slouched position, which is then coded *slouched*.

Backward lean

A rotation of the shoulders (402), also called a trunk swivel or trunk turn (Harrigan, 2005) is a movement from the waist up. In our coding this was realised without paying attention to the position of the pelvis, but with a focus on the direction towards or away from the hostess. The rotation is to be coded when the torso is turned away from the hostess, not when it is not aligned with the desk.

When the head follows the direction of the shoulders no additional annotation of the head is needed. As a consequence of the rotation one shoulder is situated in front and one in the back.

Rotation of the shoulders

A lean to the *side* (403) implies a sideways shift of the midline of the head and shoulders, which are not aligned anymore with the hips. What is more, as a consequence one shoulder is above the other.

Sideways lean

A *slouch* (404) implies “rounding” the shoulders and the upper back, that is an upper body collapse which can eventually bring the shoulders forward. This brings an impression of shrinking.

An *erect position* of the torso (405) implies an elongation of the spine. It might imply bringing the shoulders slightly backward and is often associated with an elevated “port de tête”, as a continuation of the spine.

Arms, hands and shoulders
Another body action involving the shoulders but this time with no impact on the position of the torso, is the **upward** movement of the **shoulders** (406).

Shoulders up

Another body action-position; specific to our corpus presenting a sitting person, is the placement of **elbows** on the desk (407). To code the presence of this element both elbows have to be spotted or in some cases presumed as positioned on the desk, without any possibility of coding this element unilaterally.
Fist: the only hand shape that has been defined, as important for this annotation in terms of emotional expression, was the fist. This action has to be annotated when the fingers seem to be "tightly folded". The fist is not to be coded when there is clearly some free space under the forefinger nor when the hand is relaxed.

Positions of the hand to be coded as "fist"

Position of the hand not to be confused with a fist: the fingers are relaxed (and not closed enough).

The element has to be coded whenever one or two hands are in the position. If one hand stops the tension before the second has started two actions have to recorded separately. If the tension of two hands overlaps in time, the action should be recorded as one unit of action only.

Arms crossed implies a position of the arms where each hand is held further than the other. Arms could be lying on the desk, but a simple position of support on the desk, hands clutching each other is not sufficient to score it.

Two types of **manipulators** are differentiated: functional manipulation of objects and self-manipulation. The first one is defined as movement that has a function, this function not being communicative, nor expressive, e.g. a passenger taking a ticket from the hostess who is giving it back. This action is not coded at all.
The second one codes non-functional movements, that is not helping the passenger in achieving objective goals, and not communicative. They are mostly auto-contacts, what we called "self-manipulators" (Rosenfeld, 1966), that is oriented towards the self.

Self manipulators:
Following Friesen et al. (1979) who noted that the area of the body that is being manipulated is not irrelevant in the perception and expression of these behaviours, we singled out three types of self manipulators that we coded separately:
hand self manipulator (602), face self manipulator (601), other self manipulator(600). The first one described actions that happened on the level of the hand, e.g. one finger scratching another or one hand pressing another. The second one was a touch on the level of the face, e.g. a hand moving up and down against the cheek or scratching an ear. The third coded movement involving the use of an external object, but without an aim other than the manipulation per se, e.g. drumming a finger on the desk.
For all the three self manipulators, the time of preparation is not included when this beginning of the movement is not at its target. If the passenger is scratching its cheek, the movement of the hand from the desk to the cheek is not coded as it is still undifferentiated and, to start with, may not be motivated by a self manipulation. For example a hand may be reaching towards a paper but instead of grasping it (and giving the hostess), may start playing with it.
In other words, when we code manipulators, we do not include the preparation and the retraction phases. A unit starts when the action actually starts, e.g. when a hand touches the face to scratch it (and not when the hand goes up to the face); it finishes when the hand stops touching the face. Placing a hand against another surface for no practical reason is considered a self manipulation, e.g. when placing one's head on one's arm, a hand against the cheek.

Shaking and fidgeting (603): these are to-and-fro movements (oscillations or twitching) that are fast, of small amplitude and hardly differentiated. They may affect any part of the body: e.g. shaking of the chin or of a hand.

Deep inspiration or expiration (605): we code in an undifferentiated way the inspiration and the expiration. If one follows the other without pause they are coded as one action taking place, if with a pause without movement in between, we code them as two separate movements.

Fast breathing (606): we code as such, movements of the torso that raises and lowers itself in an accelerated rate compared to the normal.

Illustrator (609)
Illustrators are movements serving to transmit visually a message. We have regrouped in this

category what Efron calls pictographs (i.e. movements drawing the shape of a referent), physiographics (the imitation of bodily actions, like running), ideographs (i.e. depicting *a direction of thought*) and deictic (Efron, 1941/1972). A deictic, or pointing, is an explicit act of drawing attention to a point of focus. By pointing is meant the directing of some part of the body, most often a finger, a hand or the head, to a person, an object or an idea (e.g. a location).

Beats/Batons (610)
Batons are movements that give a rhythm, accentuate or emphasise what is being said verbally. Efron runs a simile, comparing these movements to a conductor’s baton, beating “the tempo of...mental locomotion”. We code this category when the hands accentuate and move following the rhythm of the speech, although they do not need to be synchronous (McClave, 1994).

Appendix V
Calculation of Cohen's kappa

Cohen's Kappa (k, Cohen, 1960) is a coefficient of agreement of two raters on a nominal variable, in which two or more categories are mutually exclusive. It takes in count chance level and measures the proportion of agreement that goes beyond it.
Coefficients of 0.60 to about 0.75 indicate good, or adequate reliability; coefficients of 0.75 or higher indicate excellent reliability (e.g., Fleiss, 1981).

Formula:

$$\kappa = \frac{p_o - p_c}{1 - p_c}$$

Where p_o = the observed proportion of agreements = sum of agreements divided by total sum

p_c = the proportion of agreement expected by chance

$$p_o = \frac{\sum_i a_{ij}}{\sum_i \sum_j a_{ij}}$$

$$p_c = \frac{\sum_i \left(\sum_j a_{ij} \right) \left(\sum_j a_{ji} \right)}{\left(\sum_i \sum_j a_{ij} \right)^2}$$

a_{ji} means the value in the matrix for row I and column j

The same k calculation can be expressed in frequencies to facilitate computation:

$$\kappa = \frac{f_o - f_c}{N - f_c}$$

Appendix VI

Statistics for the total duration and mean duration of Action Units

Total duration of each action unit

Upper Face actions

AU	1	2	4	5	6	7	43
Mean	4,25	3,04	2,88	0,52	0,25	3,29	0,40
max total duration across clips	16,84	11,08	20,88	4,04	5,44	21,88	7,04

Lower Face actions

AU	9	10	11	12	13	14	15	16	17	20	23	24
Mean	0,01	1,15	0,03	0,84	0,07	2,16	0,30	0,50	1,72	0,18	1,51	0,67
max freq	0,20	8,52	0,48	11,04	1,08	17,08	2,12	14,60	8,96	1,64	14,9	5,04

Lips and jaw opening

AU	8	18	22	25	26	27	28
mean	0,08	0,07	0,03	2,84	2,91	0,05	0,32
max freq	1,72	2,52	0,76	15,07	17,96	1,20	7,48

Eyes positions

AU	61	62	63	64
mean	0,63	1,17	0,04	5,48
max freq	4,08	8,24	0,84	16,27

Head positions

AU	51	52	53	54	55	56	57	58
mean	2,06	1,95	0,94	2,81	1,10	0,49	1,87	0,30
max freq	13,52	10,80	8,52	11,88	9,32	8,88	15,96	4,72

Miscellaneous actions

AU	19	29	30	31	32	33	34	35	36	37	38
mean	0	0,14	0,00	0,06	0,03	0,00	0,23	0,00	0,00	0,01	0,00
max freq	0	2,32	0,00	2,16	1,04	0,00	8,36	0,00	0,00	0,20	0,00

Gross behaviour scores

AU	50	80	84	85
mean	4,20	0,03	0,16	0,27
max freq	26,40	0,52	4,84	2,44

Mean duration of each action unit

Upper Face actions

AU	1	2	4	5	6	7	43
Mean	1,94500623	1,80536455	1,67628626	0,3249187	0,22439024	1,70423577	0,32
max total duration across clips	9,16	11	9,44	3,88	5,44	14,28	7,04

Lower Face actions

AU	9	10	11	12	13	14	15	16	17	20	23	24
Mean	0,00	0,62	0,14	0,37	0,06	1,14	0,29	0,41	1,21	0,36	0,94	0,42
max freq	0,20	3,96	2,44	5,52	1,08	9,04	2,12	14,60	8,96	4,20	8,56	4,52

Lips and jaw opening

AU	8	18	22	25	26	27	28
mean	0,05	0,02	0,01	1,57	1,74	0,07	0,15
max freq	1,16	0,84	0,25	7,53	11,04	1,32	2,36

Eyes positions

AU	61	62	63	64
mean	0,45	0,74	0,07	2,38
max freq	3,56	5,00	0,84	15,04

Head positions

AU	51	52	53	54	55	56	57	58
mean	1,04	1,21	0,76	1,40	1,04	0,26	1,57	0,28
max freq	4,64	5,40	8,52	6,88	9,32	2,96	15,96	4,72

Miscellaneous actions

AU	19	29	30	31	32	33	34	35	36	37	38
mean	0	0,19	0,00	0,05	0,03	0,00	0,11	0,00	0,00	0,00	0,00
max freq	0	2,32	0,00	2,16	1,04	0,00	4,18	0,00	0,00	0,20	0,00

Gross behaviour scores

AU	50	80	84	85
mean	1,98	0,08	0,08	0,28
max freq	11,42	0,84	2,42	2,52

Appendix VII
Study involving a virtual character to manipulate expressions

Consigne

Nous allons vous montrer des vidéos d'un aéroport virtuel. Chaque vidéo montre un personnage virtuel différent, en train de déclarer la perte de ses bagages. Vous le voyez attendre ou interagir avec l'hôtesse virtuelle du bureau des objets trouvés (qui n'apparaît pas sur l'écran). Vous ne pouvez voir une vidéo qu'une fois. Il n'y a pas de bonnes ou mauvaises réponses et il se peut que des vidéos se ressemblent beaucoup. Essayez de répondre sans réfléchir trop longuement sur chaque question.
Il n'y a pas de son dans les vidéos.

Veuillez vous concentrer sur la vidéo avant de répondre aux questions qui suivent.

Suivant

Figure 7. Introduction guidelines for participants from the *Manipulated expression study*

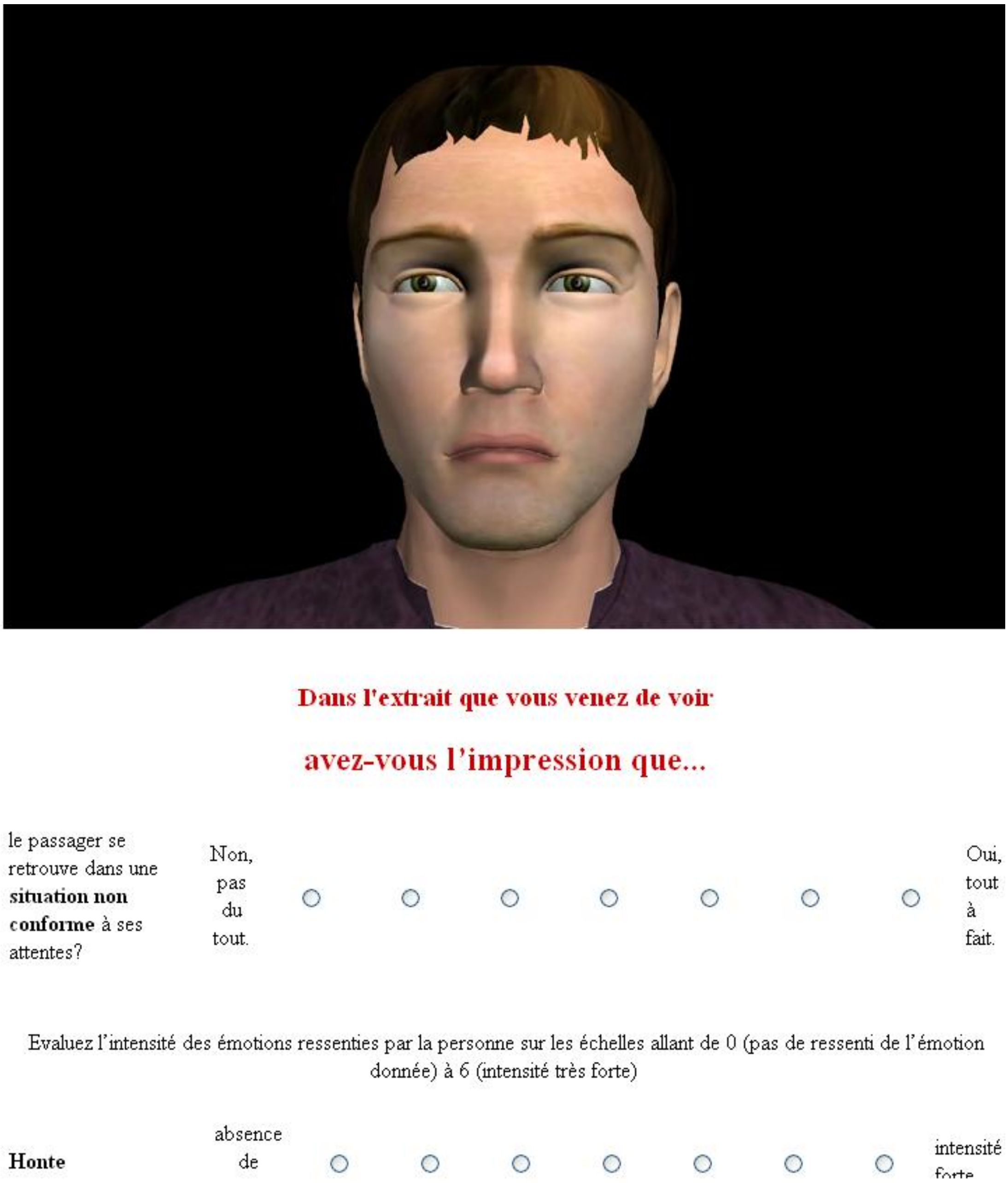


Figure 8. Sample question evaluating an appraisal outcome, provided to participants of the *Manipulated expression study* at the same time as the video to be judged (and followed by emotion labels evaluated on liker scales).