

# **The expression, experience, and social consequences of guilt: A cross-cultural study**

**Eglantine Julle-Danière**



The Leverhulme Trust

Thesis submitted in partial fulfilment of the requirements for the award of the degree of

Doctor of Philosophy of the University of Portsmouth.

**September 2019**

The face is a picture of the mind

*Cicero (46, BCE)*

## ***Abstract***

Guilt arises when someone feels s/he did something wrong; it is classified as a moral, self-conscious emotion and is one of the most social emotions that people experience throughout life. A facial expression of guilt may have evolved in humans to indicate one's willingness to make amends. In this thesis, I investigate the form and social value of facial signals associated with guilt. First, I explored the facial signals reliably associated with guilt, using a bottom-up approach, and examined the facial movements associated with the perception of guilt. I found that self-reported guilt was associated with upper lip movement and neck touching, but the perception of guilt was associated with frowning and face touching. Second, I considered the influence of contextual information and methodology on the perception of guilt. The study revealed that the context helped the interpretation of a facial expression of guilt. However, people did not reliably identify a facial signal of guilt when using emotion or action tendency labels, but did so using dimensional ratings to a certain extent. Third, I explored the tendency people have to mislabel facial displays of guilt by looking at the eye fixation patterns. I found that 'guilt' facial expressions were examined similarly to basic emotion displays and this could indicate that the emotional display of guilt was perceived as a unique mix of primary emotions. Finally, I considered the social consequences of guilt between pairs of friends. In this study, feeling guilty increased the motivation to repair wrongdoing regardless of friendship while observing guilt in others led to a punishment effect, modulated by friendship. Overall, this thesis supports the idea that guilt has an important social function and provides the first evidence that facial signals can be associated with the experience of guilt. These data build on previous research on the perception of secondary emotions and the importance of facial signals in social interactions.

## Contents list

<b>Abstract</b> .....	<b>i</b>
<b>Contents list</b> .....	<b>ii</b>
<b>Declaration</b> .....	<b>v</b>
<b>Acknowledgements</b> .....	<b>vi</b>
<b>Dissemination of research from this thesis</b> .....	<b>viii</b>
<b>List of tables</b> .....	<b>x</b>
<b>List of figures</b> .....	<b>xi</b>
<b>1. General Introduction</b> .....	<b>1</b>
1.1. Overview .....	1
1.2. How to define “emotion”? .....	4
Emotion classification: theoretical conceptualisations .....	5
Secondary emotions and culture .....	8
1.3. The communicative role of emotion.....	10
The study of facial movements.....	13
Evolution of specific expressions .....	16
1.4. Morality and moral emotions .....	17
Ontogenetic development of morality .....	17
Cooperation and punishment.....	19
Moral emotions .....	19
Shame, guilt, and embarrassment – what is the difference? .....	20
1.5. The case of guilt .....	24
What is the function of guilt? .....	24
Facial signals of guilt .....	26
<b>2. Are there non-verbal signals of guilt?.....</b>	<b>30</b>
Abstract.....	30
2.1. Background .....	31
2.2. Study 1 – Production of guilt:.....	37
Methods.....	37
Statistical Analysis.....	45
Results.....	46
Discussion .....	52
2.3. Study 2 – Judgement of Guilt .....	54
Methods.....	54
Statistical Analysis.....	61
Results.....	63
Discussion .....	66
2.4. General Discussion .....	67
<b>3. Guilt outside of context .....</b>	<b>76</b>
Abstract .....	76
3.1. Background .....	77

3.2.	Perception of guilt by a US-sample .....	78
	Participants .....	78
	Procedure .....	80
	Stimuli .....	80
	Guilt Judgements .....	81
	Compiling the dataset.....	81
	Statistical Analysis.....	82
3.3.	Results.....	83
	Descriptive .....	83
	Principal Components Analysis.....	84
	Impact of contextual information.....	85
3.4.	Discussion .....	85
<b>4.</b>	<b><i>Perception of dynamic facial expressions of 'guilt' .....</i></b>	<b>88</b>
	Abstract.....	88
4.1.	Background .....	89
4.2.	Methods.....	94
	Observers.....	94
	Demographics and acculturation measures .....	94
	Stimuli .....	95
	Procedure .....	97
	Statistical Analysis.....	98
4.3.	Results.....	99
	A. Accuracy: how good were judges at labelling guilt? .....	99
	B. Pattern of categorisation: what did judges label guilty faces as? .....	103
	C. Quantitative judgements: how did judges categorise guilt using bi-dimensional scales or emotion rulers? .....	108
4.4.	Discussion .....	113
<b>5.</b>	<b><i>Eye fixation patterns when viewing 'guilt' facial expressions.....</i></b>	<b>118</b>
	Abstract.....	118
5.1.	Background .....	119
5.2.	Methods.....	124
	Judges .....	124
	Demographics and acculturation measures .....	125
	Stimuli .....	126
	Apparatus .....	127
	Procedure .....	127
	Face Regions .....	128
	Statistical Analysis.....	129
5.3.	Results.....	130
	Task performance.....	130
	Eye-tracking data .....	134
5.4.	Discussion .....	139
<b>6.</b>	<b><i>The social function of the feeling and expression of guilt.....</i></b>	<b>143</b>

Abstract .....	143
6.1. Background .....	144
6.2. Methods .....	146
Participants .....	146
Experimental design .....	148
Procedure .....	149
Measures .....	152
Statistical Analysis.....	155
6.3. Results.....	156
Feeling guilty (Player 1).....	156
Perceiving guilt in others (Player 2) .....	160
6.4. Discussion .....	164
<b>7. General Discussion .....</b>	<b>168</b>
Overview .....	168
Main findings and implications .....	169
Facial signal of guilt as source of information .....	169
Guilt in context .....	170
Guilt and personality .....	171
Facial expression: emotional signal or predictor of future actions?.....	172
Why the confusion when looking at guilty faces? .....	172
Limitations.....	174
Future directions.....	179
Conclusions .....	182
<b>References .....</b>	<b>183</b>
<b>Appendices .....</b>	<b>195</b>
Appendix 1. Personality and Demographic Questionnaires.....	195
Big Five – “How am I in general?” .....	195
Dirty Dozen .....	197
PANAS .....	198
GASP .....	199
URCS .....	201
Acculturation Questionnaires.....	202
Appendix 2. SFEC 2015-097 Favourable opinion letter – Study 1 (Chapter 2).....	203
Appendix 3. SFEC 2016-061 Favourable with Conditions– Study 2 (Chapters 2 & 3).....	205
Appendix 4. SFEC 2016-061C Favourable Opinion, Notice of Substantial Amendment, Study 4&5 (Chapters 4&5) .....	207
Appendix 5. SFEC 2017-107 Favourable Opinion, After resubmission – Study 3 (Chapter 6).....	209
Appendix 6. UPR 16.....	211

## ***Declaration***

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

**Word count 46 000**

## ***Acknowledgements***

Unfortunately, one of the most significant things ever said about emotion may be that everyone knows what it is until they are asked to define it

*Joseph LeDoux (1996, p.23)*

Ever since Darwin, scientists agree on defining emotion as coordinated suites of behavioural, physiological, cognitive, and affective processes, summarised as automatic-motoric-cognitive states. In itself, such a definition does not bring much information. By doing a PhD, I tried to shed some light on the concept of emotion, looking mainly at behavioural expressions associated with experiences of emotion. And this PhD would have not been possible without incredible people helping me along the way.

First and foremost, I would like to thank Prof. Bridget Waller for welcoming me into her team and believing in this project since the beginning. Thank you for your continuous support and availability as my supervisor and mentor over the last 5 years. I would also like to thank Prof. Aldert Vrij for getting on board with us and providing us guidance and feedback throughout the grant application and PhD process. I am grateful to my third supervisor, Dr. Erik Gustafsson, for his helpful comments and discussion at every stage of my PhD, and Dr. Beatriz Lopez, Dr. Ed Morrison, and Dr. Ryan Fitzgerald for their comments during annual and major reviews.

I am deeply grateful to the Leverhulme Trust for funding this project and giving me a chance to pursue a dream my younger self had since she could remember: exploring the facial expression of guilt and understanding the importance of guilt in social interactions. This project would have never existed without the determination of my younger self, her curiosity and interest for facial expression, and her passion for research. We made it!



I would like to thank all of the technical staff for their help over these past four years and thank you to all of the participants! I could not have made it without you taking the time for my research.

I would like to thank everyone who helped me with data collection, reliability coding, and data analysis. More specifically, I thank Alice Durston, Sanel Sonmezani, and Hoi-Lam Jim. I am especially grateful to Jamie Whitehouse for your patience and your help through brainstorming data collection, and data analysis. Thank you for your support and advice. I would like to thank the poor souls I have shared an office with, especially in this last year; particularly Tom Lockhart and Peter Clark – thank you for putting up with me and my mood swings.

I am incredibly grateful to all the people I met throughout those years and I get to call “friends”. Thank you for reminding me there was a life outside the office, thank you for listening to my complaints, and for being there when I needed you. Claudiu, thank you for putting up with me and reminded me the PhD was not all that I was. Helene, more than anyone, we have shared this journey together and I am most thankful to have met you and for all the time we spent together, the good and the bad alike. Last, but not least, I want to thank my family, especially my parents, for believing in me, for their interest in my research, and for supporting me throughout all my adventures around the world. Thanks all, I would not have made it without you.

## ***Dissemination of research from this thesis***

Publications, abstracts, presentations, and posters

### **Chapter 1**

- Waller, B.M., Julle-Danière, E., & Micheletta, J. (under review). Measuring the evolution of facial expression using multi-species FACS. *Neuroscience and Biobehavioural Reviews*

### **Chapter 2:**

- Julle-Danière, E., Whitehouse, J., Vrij, A., Gustafsson, E., & Waller, B.M. (under review). A facial signal of guilt in humans. *PLOS One*
- Julle-Danière, E., Whitehouse, J., Vrij, A., Gustafsson, E., & Waller, B.M. 2016. The facial expression of guilt: Myth or reality? Oral Presentation at the Consortium of European Research on Emotion (CERE; July 4-7<sup>th</sup>)
- Julle-Danière, E., Whitehouse, J., Vrij, A., Gustafsson, E., & Waller, B.M. 2017. The facial expression of guilt: production and cultural differences. Paper presented at the International Society for Research on Emotion meeting (IRSE; July 26-29<sup>th</sup>)
- Julle-Danière, E., Whitehouse, J., Vrij, A., Gustafsson, E., & Waller, B.M. 2018. The facial expression of guilt: production and perception. Oral Presentation at the Consortium for European Research on Emotion meeting (CERE; April 4-5<sup>th</sup>)

## Chapter 6:

- Julle-Danière, E., Whitehouse, J., Vrij, A., Gustafsson, E., & Waller, B.M. (under review). Friendship modulates the punishment function of guilt. *Proceedings of the Royal Society B*
- Julle-Danière, E., Whitehouse, J., Vrij, A., Gustafsson, E., & Waller, B.M. 2019 Friendship modulates the punishment function of guilt. Oral Presentation at the International Society for Research on Emotion meeting (IRSE; July 10-13<sup>th</sup>)

## *List of tables*

TABLE 2.1. PARTICIPANTS DEMOGRAPHIC INFORMATION (STUDY 1).....	38
TABLE 2.2. LIST OF ACTION UNITS (AUS), MOVEMENTS, AND MISCELLANEOUS ACTIONS CODED FOR	43
TABLE 2.3. SELF-REPORTED EMOTIONS BEFORE AND AFTER GUILT INDUCTION .....	48
TABLE 2.4. FACTOR LOADINGS OF THE PRINCIPAL COMPONENT ANALYSIS (PCA) – GUILT CONDITION	50
TABLE 2.5. FACTOR LOADINGS OF THE PRINCIPAL COMPONENT ANALYSIS (PCA) – CONTROL CONDITION .....	51
TABLE 2.6. PARTICIPANTS DEMOGRAPHIC INFORMATION (STUDY2).....	55
TABLE 2.7. DESCRIPTIVE OVERVIEW OF GUILT PINPOINTS IDENTIFIED BY JUDGES .....	64
TABLE 3.1. PARTICIPANTS DEMOGRAPHIC INFORMATION.....	79
TABLE 3.2. DESCRIPTIVE OVERVIEW OF GUILT PINPOINTS IDENTIFIED BY JUDGES .....	83
TABLE 3.3. FACTOR LOADINGS OF THE PRINCIPAL COMPONENT ANALYSIS (PCA) .....	84
TABLE 4.1. EMOTION WORDS AND ACTION TENDENCIES USED IN THE FORCED-CHOICE TASKS .....	97
TABLE 4.2. PROPORTION OF JUDGES WHO MATCHED THE PREDICTED EXPRESSION TO THE CORRESPONDING EMOTION OR ACTION TENDENCY LABEL.....	101
TABLE 4.3. PROPORTION OF JUDGES WHO MATCHED THE PREDICTED EXPRESSION TO THE CORRESPONDING EMOTION OR ACTION TENDENCY LABEL BASED ON THE CATEGORISATION METHOD USED .....	102
TABLE 4.4. PROPORTION OF JUDGES WHO MATCHED THE PREDICTED EXPRESSION TO THE CORRESPONDING LABEL (EMOTION OR ACTION TENDENCY) DEPENDING ON THE CULTURE OF THE JUDGE .....	104
TABLE 4.5. PROPORTION OF JUDGES MATCHING A FACIAL EXPRESSION TO A LABEL DEPENDING ON THE CULTURE OF THE JUDGE.....	105
TABLE 4.6. PROPORTION OF JUDGES MATCHING A FACIAL EXPRESSION TO AN EMOTION OR ACTION TENDENCY LABEL.....	107
TABLE 4.7. MEAN AND STANDARD DEVIATION OF AROUSAL AND PLEASANTNESS RATINGS FOR EACH EMOTION TYPE.....	110
TABLE 4.8. MEAN AND STANDARD DEVIATION OF EMOTIONAL RATINGS FOR EACH EMOTION TYPE	112
TABLE 5.1. PARTICIPANTS DEMOGRAPHIC INFORMATION .....	125
TABLE 5.2. CROSS TABLE RESULTS SHOWING CHOICES MADE FOR EACH FACIAL EXPRESSION.....	133
TABLE 5.3. LINEAR MIXED MODEL FOR 4 FACE REGIONS LOOKING AT CULTURAL CONGRUENCY BETWEEN JUDGES AND VIDEOS.....	135
TABLE 6.1. PARTICIPANTS DEMOGRAPHIC INFORMATION .....	147
TABLE 6.2. LIST OF ACTION UNITS (AUS), MOVEMENTS, AND MISCELLANEOUS ACTIONS CODED FOR	153
TABLE 6.3. MEAN AND STANDARD DEVIATION OF THE SELF-REPORTED EMOTIONS THROUGHOUT THE STUDY .....	157
TABLE 6.4. ANNOYANCE TOWARDS SELF AND PARTNER MEASURED AFTER THE COOPERATIVE GAME	159

## *List of figures*

<b>FIGURE 1.1. DIFFERENTIATE SHAME, GUILT, REMORSE AND REGRET.</b> .....	23
<b>FIGURE 2.1. GENERAL PROCEDURE.</b> .....	39
<b>FIGURE 2.2. IMAGES OF GUILTY EXPRESSIONS TAKEN FROM VIDEOS.</b> .....	52
<b>FIGURE 2.3. JUDGEMENT RATINGS.</b> .....	59
<b>FIGURE 2.4. REPRESENTATION OF THE TEMPORAL PRODUCTION OF FACIAL MOVEMENTS.</b> .....	65
<b>FIGURE 4.1. GENERAL PROCEDURE.</b> .....	98
<b>FIGURE 4.2. ACCURACY SCORES PRESENTED FOR EACH FACIAL EXPRESSION IN EACH CONDITION.</b> .....	101
<b>FIGURE 4.3. ACCURACY SCORES PRESENTED FOR EACH FACIAL EXPRESSION FOR EACH METHOD USED.</b>	102
<b>FIGURE 4.4. ACCURACY SCORES PRESENTED FOR EACH FACIAL EXPRESSION FOR EACH CULTURAL GROUP.</b> .....	103
<b>FIGURE 4.5. SCATTER PLOT SHOWING THE RATINGS FOR EACH FACIAL EXPRESSION IN THE BI- DIMENSIONAL SPACE</b> .....	109
<b>FIGURE 4.6. EMOTION RATINGS FOR EACH FACIAL EXPRESSION (IN PERCENT)</b> .....	111
<b>FIGURE 5.1. ILLUSTRATION OF 10 FACIAL REGIONS OF INTEREST (ROIS).</b> .....	128
<b>FIGURE 5.2. ACCURACY SCORES PRESENTED FOR EACH FACIAL EXPRESSION IN EACH CULTURAL GROUP</b> .....	132
<b>FIGURE 5.3. A. AVERAGE FREQUENCY OF FIXATION FOR EACH CONGRUENT GROUP; B. AVERAGE FREQUENCY OF FIXATION TO EACH FACE REGION.</b> .....	136
<b>FIGURE 5.4. AVERAGE FREQUENCY OF FIXATION TO EACH FACE REGION DEPENDING ON THE CULTURAL GROUP</b> .....	136
<b>FIGURE 5.5. AVERAGE FREQUENCY OF FIXATION TO EACH FACE REGION DEPENDING ON THE FACIAL EXPRESSION PRESENTED.</b> .....	138
<b>FIGURE 6.1. GENERAL DESIGN.</b> .....	149
<b>FIGURE 6.2. COOPERATIVE GAME.</b> .....	150
<b>FIGURE 6.3. THE INTERACTION BETWEEN JUDGED GUILT AND FRIENDSHIP.</b> .....	162

## **1. General Introduction**

### **1.1. Overview**

Guilt is an emotional and cognitive experience arising when someone feels that they did something wrong. It could have evolved due to its potential adaptive function, within social interaction, of stimulating pro-social behaviours towards and from others. Guilt is classified as a moral, self-conscious emotion and is one of the most social, other-oriented emotions that people experience throughout life. People often state that they can tell when someone is expressing remorse or feeling guilty but, with some notable exceptions, to date a facial expression of guilt had not been identified scientifically. Whether emotions (and which emotions) are associated with universally produced and recognised facial expressions is debated. Due to its later ontogenetic development, the experience of guilt is probably idiosyncratic and context-dependent, as well as influenced by cultures and social norms. A facial expression of guilt may have evolved in humans due to the value in indicating one's willingness to make amends. Previous research has tried to identify a recognisable set of facial movements associated with the experience of guilt, but with no success. The aim of this thesis is to explore the emotion of guilt from three intertwined perspectives: the behavioural expression of guilt, the experience of guilt, and the responses to guilt from observers. I investigate this by combining the induction of genuine feelings of guilt with judgement experiments to broaden our understanding of both production and perception of facial movements associated with guilt.

There is still much debate regarding the definition of emotion and how emotional facial displays should be interpreted. To introduce this thesis, in this first chapter (General Introduction), I present the different schools of thoughts and highlight the definitions chosen for this investigation. First, I introduce the concept of emotion and review the successive classification theories that have been developed in the last 50 years, emphasising the differences

between various moral emotions, namely shame, embarrassment, and guilt. I also discuss the communicative value of the face, considering how it supports the rapid nonverbal transmission of socially relevant information, such as emotional states. Finally, I review the current state of the art regarding guilt and highlight the missing pieces of the puzzle.

In the first empirical chapter, Chapter 2, I investigate variation in both the production and perception of the specific facial movements associated with guilt. I examine the production and perception of spontaneous facial expressions using a bottom-up approach to identify dynamic patterns in facial behaviour. Firstly, using an experimental induction approach and an extensive dynamic facial movement coding system, I have been able to identify patterns of movements that were associated with the subjective feelings of guilt. Secondly, by presenting videos taken during the experience of guilt to new participants, I begin to understand how guilt is perceived by social partners and identified patterns of facial movements people associated with feelings of guilt. These findings suggest that there is a basic non-verbal pattern identified as guilt by observers, but that this is not highly correlated with felt guilt in the actor.

In Chapter 3, I examine the impact of contextual information on the ability to detect guilt on a face. Secondary emotions have been shown to be context and culture-dependent in terms of subjective experience; I test here if this also applies to the decoding of expressions associated with the experience of guilt. To test this, I replicate the judgement study from Chapter 2 with a new sample of participants and I randomise the presentation of contextual information. I measure the accuracy level of judges to identify guilt on a face and look at the impact of context on this accuracy. I replicate all my previous findings regarding the perception of guilt and I also show that a facial signal of guilt is better interpreted when the situational context is provided alongside the facial expression. It supports the trend for ecologically valid judgment research as a facial display is rarely occurring without context in real life.

Emotion judgement studies are typically conducted using emotion words. However, a growing literature shows that facial displays can be understood in terms of action tendencies. In Chapter 4, I compare categorisation methodology to shed light on how facial displays of guilt are best interpreted: in terms of emotion words, action tendencies, or using a bi-dimensional space. Firstly, in the absence of contextual information, people struggle to identify facial displays of guilt, whereas this does not impair their ability to identify facial expressions associated with primary emotions, regardless of the labels used (emotion words or action tendencies). Secondly, guilty displays are reliably differentiated from other expressions when using a bi-dimensional space. Therefore, I build on previous literature, showing that secondary emotions may be reliably interpreted by observers, but not necessarily using categorical, discrete labels and so might not necessarily be conceptualised as discrete phenomena.

In Chapter 5, I investigate the scanning patterns associated with the recognition of guilty expressions, comparing ocular movements in two culturally distinct populations. If people stare preferentially at a specific area of the face, it could help explain the mis-categorisations observed in previous chapters. Firstly, I find that 'guilt' facial expressions are examined in a similar manner to displays of the basic emotions, as observers display specific fixation patterns associated with the process of faces displaying emotional content. Secondly, I demonstrate that observers exhibit scanning patterns that overlap with some of the previously identified patterns used when looking at facial signals of fear. Finally, I find cultural differences in the viewing patterns of guilt displays. Therefore, I replicate previous findings showing that different cultures process faces differently and that it is possible that observers perceive emotional displays of guilt in terms of valence and arousal, placing the emotional experience on a bi-dimensional space, rather than in terms of a specific emotion.

In Chapter 6, I look at the social aspect of guilt, investigating the behavioural outcomes for both the wrongdoer and the victim, and how these outcomes are modulated by friendship. If



guilt truly has such a strong social component, there should be differences in the expression of guilt or reaction to guilt depending on the quality of the relationship. Here, I examine the consequences of guilt in an ecologically valid setting by observing a pair of interacting individuals. Firstly, I replicate previous findings showing the repair function of guilt and I find that guilty people are motivated to repair their wrongdoing, regardless of friendship quality. Secondly, I demonstrate that observing guilt in a social partner lead to a punishment effect, and this effect is modulated by the quality of the friendship between social partners. Indeed, victims of wrongdoing punish their close friends who appeared guilty more than acquaintances. Therefore, I suggest that guilt has a punishment function modulated by friendship and that the relationship context is crucial to the evolutionary relevance of moral emotions.

In the seventh and final chapter of this thesis, I will summarise all the mentioned studies, highlighting the key findings as well as theoretical and practical implications of the results. Overall, this thesis supports the idea that guilt has an important social function and provides the first evidence that facial signals can be associated with the experience of guilt. This data builds on previous research on the perception of secondary emotion and the importance of facial signals in social interactions.

## 1.2. How to define “emotion”?

The definition of emotion as we know it today comes from the 19th century and is often attributed to Darwin, who first tried to understand the reason why emotions have evolved, i.e. what function they serve, before trying to define emotion. Darwin originally attributed emotion a warning function, preparing the body, or the organism, to respond adaptively and rapidly to environmentally recurring stimuli (Darwin, 1872). So emotion, or feeling, is an internal state resulting from physiological changes, triggered by our social and physical environment,

triggering behavioural, cognitive, and affective processes (Shariff & Tracy, 2011). What we define as 'emotion' is then a physiological response. Based on this Darwinian definition, it seems reasonable to think emotion is not a human-specific concept; any animal living in a given habitat will face environmentally recurring stimuli that could affect the animal's fitness (i.e. its survival ability and reproductive success). Thus, when taking an evolutionary approach to the study of emotion, which would help understand the origin of and reason for human emotion, other species should also be considered. The difference between human and animal emotion might be in degree, not in nature (de Waal, 2011). Human societies are one of the most complex forms of social and cultural system on earth (Boyd & Richerson, 2009; Hill & Dunbar, 2003). Social systems exist however in a whole range of species (Whitehead, 2008), as some sort of organisation between individuals is necessary for survival. The distinction is thus in the complexity of the given organisational system, impacting the complexity of a social environment, which in turn influences the physiological responses displayed in an individual. As our social environment developed more complexity, the triggered changes became in turn more complex, which led to the experience of complex emotional situations and the need to classify and categorise more emotion.

### **Emotion classification: theoretical conceptualisations**

Depending on their ontogenetic origins, their purpose, or their objects (self vs other), emotions can be categorised differently. Emotions that appear early in life are distinguished from those that develop gradually during childhood. These two types of emotion are often called primary and secondary emotions (Darwin, 1872). So-called primary emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise) are said to be biologically based, shared with other animals, pan-culturally expressed, and easily studied without verbal communication (emotional state inferred from facial expressions) (emotional state inferred from facial expressions; Tracy & Robins, 2004). Those emotions appear first in life, supposedly emerging

between 1 to 9 months after birth (Campos, Barrett, Lamb, Goldsmith, & Stenberg, 1983). In contrast, secondary emotions (e.g., guilt, shame, pride, contempt) are thought to be acquired later in life and to start developing gradually in children, from 18 months onward (Lewis, Sullivan, Stanger, & Weiss, 1989; Tangney, 1999). This later ontogeny led scientists to explore the possible influence of nurture (i.e., family environment, culture, social context) on the development of secondary emotions. Researchers agreed on this first dichotomy between primary and secondary emotions. However, the scientific community do not agree on the best theory to provide a detailed account of all the emotions experienced by individuals throughout their life (Cowen, Sauter, Tracy, & Keltner, 2019; Crivelli & Fridlund, 2019; Ekman & Cordaro, 2011; Fridlund, 2017).

The classic and largely dominant view - the Basic Emotion Theory (BET) - is that primary basic emotions (Ekman & Cordaro, 2011) are considered innate to all human populations and universally expressed (Brown, 1991; Ekman & Friesen, 1969). The classification of surprise and disgust as primary emotion was debated for a long time (Panksepp, 2007; Tracy & Randles, 2011; Turner, 2000; p.68) but most researchers would concede that they belong in this category (Ekman & Friesen, 1969; Ekman & Keltner, 1970; Field, Woodson, Greenberg, & Cohen, 1982; Levenson, 2011; Ortony & Turner, 1990). Those six emotions seem to be the basic emotions that allow infants to express themselves and communicate their basic, vital, needs to their relatives (Ganchrow, Steiner, & Daher, 1983; Steiner, 1979). The BET supports the idea that emotion can be separated into specific categories (Ekman & Cordaro, 2011), each category differing in their appraisal and behavioural outcomes (e.g., facial expressions). As a result of this, researchers have been looking, so far unsuccessfully, for emotional “signatures” or “fingerprints”, i.e., physical and physiological changes, resulting from distinctive patterns of the autonomic nervous system activity, that reliably indicate the presence of one given emotion (Barrett, Mesquita, Ochsner, & Gross, 2007; Gross & Barrett, 2011; for review, see Siegel et al., 2018).

Another approach takes a dimensional perspective; emotion could be interpreted in terms of bipolar dimensions rather than discrete names and classes (Daly, Lancee, & Polivy, 1983; Russell, 1989; Watson & Tellegen, 1985). A feeling can be more or less pleasant, or vary in arousal; by rating emotion in terms of degrees (pleasure, arousal), broad clusters can be created providing a general representation of different types of emotion (e.g., pleasant, low arousal; mid-pleasant, high arousal; unpleasant, high arousal). Valence and arousal are arguably at the core of all emotional experiences (Cowen & Keltner, 2017). This bi-dimensional conception might then provide an account of what is universal in how emotion is perceived, whereas the categorical view might be tied to cultural and semantic limitations (Barrett et al., 2007; Cowen & Keltner, 2017).

An alternative view relies on the interpretation of facial expressions as a response to a type of situation or as part of an instrumental action (Crivelli & Fridlund, 2019; Fridlund, 1994, 2017). This view distanced itself from the Darwinian definition of emotion, as emotion is not seen purely as a physiological response, but instead as a social tool (Darwin, 1872; Frijda, 1986). When trying to describe one's current feeling, it is quite frequent to start describing similar situations in which comparable states were experienced ("I feel like when..."). This approach seems more intuitive, quick, automatic, and requiring less effort than when looking for the precise words pinpointing accurately the emotional state. The limitation is that one situation might not trigger the same response in everybody, and as such, this method might not be as generalisable as the previous methods mentioned.

Those different theories, however, do not account for the full emotional experience someone will have throughout life. Indeed, classic categorisation theories rely on the ability of a person to know exactly how they feel at a given time and to be able to describe accurately their emotion using single words (Barrett et al., 2007; Ekman & Cordaro, 2011) or social situations (Crivelli, Russell, Jarillo, & Fernández-Dols, 2016, 2017; Fridlund, 1994, 2017). The use

of bi-dimensional space removes those constraints but loses at the same time the precise distinction between affective experiences (Barrett, 2006; Cowen & Keltner, 2017). Recent research tried to move away from those classical views to create a high-dimensional taxonomy of affective experiences (using up to 13 different scales, such as valence, dominance, and certainty; Cowen et al., 2019). In this thesis, I compared multiple theories to grasp how guilt was expressed and perceived in social situations.

### **Secondary emotions and culture**

Secondary emotions, such as guilt, are often perceived as deriving from primary emotions, being a mix between primary emotions arising at the same time (Russell & Fernández-Dols, 1997; The Facial Expression Program point 5, Turner, 2000, p.72). For example, someone could feel angry and disgusted in a situation, resulting in textbook contempt; whether someone can *de facto* recognise contempt will depend on the balance between anger and disgust felt at the time. People could report feeling angry, as it is the main emotion they perceived and identified, disgust having a lower arousal. Some people might be able to recognise contempt if both anger and disgust are expressed at similar intensities. Secondary emotions are also different shades of a primary emotion. For example, a situation can stimulate anger, but it can also go further and provoke rage or fury, or only provoke mellow annoyance. Rage and fury are defined as “intense anger”, their arousal is higher than anger, but annoyance is also a shade of anger, defined as “slight anger”, and is milder than rage (lower arousal); the root to both emotions remains the same (Plutchik, 1980).

A common analogy is to associate each primary emotion with a primary pigment colour (red, yellow, and blue); secondary emotions are thus a mix between two primary emotions, i.e. green results from associating blue and yellow (Mohr, Jonauskaitė, Dan-Glauser, Uusküla, & Dael, 2018). Based on the quantity of each of the primary pigment you put into the mix, the

more intense the green colour. You could also decide to add water to a primary pigment, resulting in a less vivid colour, a fainter emotion harder to identify. Mixed feelings can be tricky to pinpoint; it requires a good knowledge of yourself, the ability to listen to your body, and being able to recognise the signs of different emotions. Imagine you are looking at a nice peacock blue painting; depending on your knowledge on colours, your sensibility to different pigments, the surrounding light, and your tastes, you will label it as peacock blue, petrol, turquoise, or even dark aqua. If all those shades of colours are quite close, they are different to the shrewd eye as the exact intensity of each pigment varies in each of those shades. The same happens with secondary emotions (Barrett, 2017a; Plutchik, 1980; Russell & Fernández-Dols, 1997). Philosophers and scientists tried to define a more objective classification system of secondary emotions, based on their functions (i.e. social or moral emotions), their objects (i.e. self- vs other-oriented), and their valence (i.e. positive or negative; Thamm, 2006). The system might look complex, as quite often an emotion will have multiple labels (i.e. pride is a positive moral self-conscious emotion), but it allows flexibility in the system and it is a more instinctive split.

Considering that secondary emotions could be subject to cultural variations (Matsumoto, Yoo, Hirayama, & Petrova, 2005), researchers started investigating what happens when a person moves between cultures i.e. will the set of secondary emotions experienced by someone in one culture remain if that person changes culture? Whether people experience a specific emotion seems to be related to their cultural norms (Mesquita & Leu, 2007). A specific emotion could be experienced more in one culture (i.e., anger in European countries) and people in different cultures can experience emotions differently (i.e., anger associated with a feeling of control in European countries but a feeling of guilt in East-Asian countries), and these processes do not seem genetically fixed. Indeed, previous research has shown an ability for people changing cultural environment to 'adapt' to their new culture by experiencing emotional acculturation and beginning to experience emotion according to their new cultural standards

(De Leersnyder, 2017). The cultural impact on emotional experience is thus not limited to early childhood exposure, but can also influence the experience of adults changing culture.

### 1.3. The communicative role of emotion

Communication is the cornerstone of successful relationships, either professional or private, as the way to share information and agree on future actions. More than just communicating on facts, it is important to indicate emotional states to our relatives to ensure long-lasting connections, using verbal and non-verbal cues and signals. This is achieved primarily by the exchange of a set of social signals, such as facial expressions and body postures (Darwin, 1872; Elenbaas, Beaupré, Lévesque, & Hess, 2007; Marsh, Elenbaas, & Ambady, 2003). Indeed, if some physiological functions have been retained throughout evolution, the main purpose of emotional expressions in contemporary human life may have more to do with the rapid nonverbal transmission of socially relevant information (Shariff & Tracy, 2011). As such, facial displays can be used to convey messages regarding internal states, emphasizing or contradicting the verbal speech, or even occurring in its absence (Horstmann, 2003). Facial expressions, more than any other nonverbal signals, are thought to be the physical representations of our emotions (Ekman & Friesen, 1978; Izard, 1997).

A signal is a behaviour that has evolved with the specific purpose of being communicative (i.e. pointing gesture to direct attention to a specific object/location; Dehaene, Mercier, & Scott-Phillips, 2013; Krebs & Davies, 1993). A cue on the other hand gives information accidentally, as a repercussion or by-product of something serving another adaptive purpose (i.e. chewing indicates someone is eating but the primary function of chewing is breaking down the food vs. being an indication that someone is eating; Dehaene et al., 2013; Krebs & Davies,

1993). It has been hypothesised that facial expressions of emotion started as cues, the by-product of the physiological changes happening internally, and evolved in both form and function into signals (Shariff & Tracy, 2011).

Whether emotions (and which emotions) are associated with universally produced and recognised facial expressions, however, is strongly debated. Ekman and colleagues (Ekman & Friesen, 1969) showed that non-European cultures were able to identify and label facial displays of emotion in the same way as Europeans, concluding that the message relayed by facial expression must be universal and so the result of natural selection (Ekman & Keltner, 1970). It is often claimed that secondary emotions do not have a prototypical universal expression (Izard, 1994) and that the experience of the secondary emotions themselves can differ significantly between cultures (Bedford & Hwang, 2003; Matsumoto, 1989).

However, the evidence for strictly universal primary emotions and culturally variable secondary emotions has been challenged and the topic is subject to debate. Over the last 50 years, a body of research has demonstrated cultural differences in the expression (Chen & Jack, 2017) and perception of facial expression of primary emotions (Elfenbein et al., 2007; Jack, Caldara, & Schyns, 2012), often comparing European and Asian cultures (see Crivelli & Fridlund, 2019 for review on BET shortcomings). The universality in the perception of basic emotion has also been found to be method-bound (Fridlund, 2017; Gendron, 2017). One study used “prototypical” expressions of emotion in a design comparing the BET classical approach (facial expressions are perceived as emotional signals; Ekman & Friesen, 1969) to the behavioural ecology approach, where facial expressions are instead seen as functional communicative signals (Behavioural Ecological View - BECV; Crivelli & Fridlund, 2019; Fridlund, 2017). Using a cross-cultural perspective, the authors argued that faces convey behavioural intentions in addition to emotional messages (Crivelli et al., 2016). Another study used mathematical modelling of dynamic face movement patterns to look at the interpretation of various facial



expressions. They found four fundamental patterns, instead of the six “universal” facial expressions, and argued for a symbiosis of biology and culture with culture-specific accents (Chen & Jack, 2017; Jack, Sun, Delis, Garrod, & Schyns, 2016). Similarly, there is some evidence that some secondary emotions are associated with recognisable facial movements, challenging the distinction between the two groups of emotion. For example, shame and embarrassment are associated with recognisable patterns of facial movements, although they are not as predictable and prototypical as the primary emotions (Ekman & Keltner, 1970; Keltner & Buswell, 1996).

It is possible, therefore, that secondary emotions, such as guilt, have common facial expressions that have been overlooked. Some cultural differences (or accents; Marsh et al., 2003) might also be expected, which would not necessarily contradict an evolutionary view. Different cultures may rely on different facial clues, which are functional in a given culture and for a given emotion, and not necessarily in a different culture (Jack, Garrod, Yu, Caldara, & Schyns, 2012; Keller & Otto, 2009). Whether people experience a specific emotion seems related to their cultural norms (Mesquita & Leu, 2007). A given emotion could be experienced more so in one culture (i.e., anger in European countries). People in different cultures can also experience emotions differently (i.e., anger associated with a feeling of control in European countries but a feeling of guilt in East-Asian countries), and these processes do not seem genetically fixed. Indeed, previous research has shown the impact of culture and environment (i.e. nurture) on the production of facial expression was stronger than the impact of ethnicity (i.e. nature; Camras, Bakeman, Chen, Norris, & Cain, 2006). For instance, adopted Chinese girls displayed a level of expressivity significantly higher than Mainland Chinese girls and Chinese American girls, revealing a strong impact of culture and family attitudes on facial expressivity. So, possible innate differences between ethnicity can be superseded in favour of cultural influences (Camras et al., 2006). Some research has been conducted on the acculturation in emotional expression (for review see Cordaro et al., 2018), i.e. does the expression or perception of a given emotion

change over time? Previous research demonstrated that individuals from different cultures used different cues to identify emotional facial expressions (Jack, Caldara, et al., 2012; Yan, Young, & Andrews, 2017) or produce different facial expressions in the same situation (De Leersnyder & Mesquita, 2015). However, it has also been shown that time spent in a culture improved the recognition of a culturally specific facial expression (Elfenbein & Ambady, 2003), suggesting that people can become acculturated to different styles of emotional recognition.

### **The study of facial movements<sup>1</sup>**

The investigation of human facial non-verbal communication has been greatly facilitated and standardised by the development of the Facial Action Coding System (FACS; Ekman & Friesen, 1978; Ekman, Friesen, & Hager, 2002). Prior to this, the human facial expression field was reliant on subjective methods and did not have a systematic way to assess the musculature components of facial expression (and thus help determine homology). Duchenne de Boulogne (1862/1990), however, was the first scientist to pursue a strongly anatomical approach to human facial expression and conducted a series of electrical stimulation studies to try and link facial muscles to specific expressions. Duchenne wanted to understand how facial landmarks shaped facial expressions, and map the connection between the contraction of individual facial muscles and observable facial displays. Building on this seminal work, Hjortsjo (1970) was the first to try and use an understanding of the relationship between facial muscle contraction and facial movements to develop a usable coding scheme for research. Hjortsjo (1970) attempted to identify the smallest independent units of muscle movements in the face, and use numbered codes to refer to their appearance changes on the face. This approach set the scene for FACS development.

---

<sup>1</sup> section published in Waller, B.M, Julle-Danière, E, & Micheletta, J. (under review). Measuring the evolution of facial expression using multi-species FACS. *Neuroscience & Biobehavioral Reviews*

Following directly from the anatomical work of Duchenne (1862/1990) and Hjortsjo (1970), FACS identifies the appearance changes related to facial movements and aims to identify individual muscle contractions, focussing not on the expression of emotions but on the production of spontaneous facial movements. For example, FACS is able to compare facial expressions objectively across individuals regardless of the inherent variability in the surface morphology of faces, e.g., bone structure, fatty deposits, skin texture, and individual muscle variations (Waller, Bard, Vick, & Smith Pasqualini, 2007; Waller, Parr, Gothard, Burrows, & Fuglevand, 2008). There is some debate, however, about the underlying assumptions of FACS. Some studies suggest that facial musculature is not consistent between individuals. Muscles sometimes differ in term of presence, size and symmetry (McAlister, Harkness, & Nicoll, 1998; Waller et al., 2008) as well as in fatty deposit and in neural supply (Ekman, 1980). Neuropsychological studies have also shown a greater involvement of the left half of the face in the expression of facial expressions, leading to asymmetrical displays of emotion (Borod, Haywood, & Koff, 1997). Moreover, some people have greater facial flexibility and/or control over their facial muscles, allowing for the production or suppression of more facial movements (Cole, Zahn-Waxler, Fox, Usher, & Welsh, 1996). However, the success with which FACS can be applied suggests that these issues do not affect the application of FACS in any great depth.

FACS uses numbers to refer to the appearance changes associated with 33 facial muscle contractions (Action Units [AUs]) and 25 more general head/eye movements (Action Descriptors [ADs]). Most AUs refer to the contraction of single muscles, but some muscles always co-occur, or are capable of producing different movements. Thus, the correspondence between facial muscles and movements is not always direct. It presents each AU in terms of underlying musculature (location and direction of action), appearance changes (multiple cues for identifying AUs), reference for AUs (subtle differences between AU combinations), how to do the AU (voluntary production of AU in isolation), and intensity scoring for the AU (criteria for

coding decisions). Because the system is based on the premise that muscles vary only slightly within species (although this is largely an untested assumption), FACS can compare facial movements regardless of superficial individual differences in other aspects of facial anatomy, such as hair covering, the facial colouration, bone structure, etc. This latter characteristic also makes FACS ideal for modification across species. The development of such a comprehensive coding system with the common language of AUs, with numerical codes and neutral labelling, has enabled researchers across a wide variety of sub-disciplines, often with diverging theoretical positions, to communicate and evaluate findings using a common language (Ekman & Rosenberg, 1997). Moreover, the FACS method is particularly well suited to comparative studies (cross-cultural and cross-species) as it provides clear descriptions for the identification of each AU, listing various appearance changes (movement of facial landmarks, changes to the shape of facial features) that can be directly compared. FACS has even been used to try and describe the facial movements depicted on archaeological material culture (Samson & Waller, 2010). In sum, FACS has become the most widely used coding system in facial expression research and requires training and certification to be used.

Since its creation, FACS has been used extensively in research (the original 1978 manual is cited by over 1025 articles; Google Scholar search in March 2019) and has been adapted for the study of facial displays of primary emotions (EMFACS; Friesen & Ekman, 1983) and for infants (BabyFACS; Oster, 2006), allowing researchers to investigate the facial expression of pre-linguistic infants (e.g., Longfier et al., 2016; Soussignan et al., 2018). FACS offers great flexibility for use in scientific research and is largely a-theoretical in the sense that it is purely a methodological tool. FACS can be used to code occurrences of AU/ADs (i.e. frequencies of each AU/AD: Galati, Sini, Schmidt, & Tinti, 2003), duration of AU/ADs (i.e. for how long was each AU/AD produced: Reed, Zeglen, & Schmidt, 2012), and intensities, using either the full 5-point scale presented in the manual or any suited adaptation (i.e., small vs big intensity of movement; small vs intermediate vs max intensity). Moreover, FACS can also be used to code for specific

AUs of interest, identified by the researchers based on previous literature or conceptualisation (K. L. Schmidt, Bhattacharya, & Denlinger, 2009) or to code for all facial movements produced, using a bottom-up approach (Julle-Danière, Whitehouse, Vrij, Gustafsson, & Waller, under review).

As such, FACS has been used extensively in the study of facial expression of emotion (Keltner & Buswell, 1996) and helps standardise data and stimuli for cross-cultural studies (Crivelli et al., 2017). This methodology also allowed for the creation of computerised stimuli or avatars posing genuine facial expressions (Jack et al., 2016), or for the development of automated analysis of facial movements (Lien, Kanade, Cohn, & Li, 1998). More specifically, the development of EMFACS has been critical for developing automated coding systems (Lien, Kanade, Cohn, & Li, 2000; McDuff et al., 2016). However, it is important to note that FACS in itself is atheoretical, and how it is applied matters. The AU profiles of the prototypes of the six basic emotions published along with the human FACS (FACS Investigators Guide: Ekman et al, 2002a/b) are used extensively as stimuli in research, but the typicality and relevance of these expressions across cultures has been questioned (Jack, 2013 - visual cognition; Feldman-Barrett, 2019). Instead, we advocate using FACS simply as a tool to measure the production of facial expression in detail and objectively.

### **Evolution of specific expressions**

Darwin argued that some facial expressions were the results of selective processes and thus served a definite and specific communicative function to reveal emotional states (Darwin, 1872). This view was then adopted by the BET researchers, positing facial expressions found precursors in the behaviours of other mammals in comparable contexts (Ekman & Davidson, 1994; Matsumoto, Keltner, Shiota, O'Sullivan, & Frank, 2008; Shariff & Tracy, 2011). The BECV (Crivelli & Fridlund, 2019; Fridlund, 2017) argued that facial expressions did not evolve to express

emotions but social intentions or motives. If they disagree on the evolutionary reason of facial expressions, both theories agree that facial signals are the products of evolution and thus serve a role in social interactions (Keltner, Sauter, Tracy, & Cowen, 2019). Moreover, a wide array of research has demonstrated the similarity in facial displays exhibited throughout the primate order (Davila-Ross, Jesus, Osborne, & Bard, 2015; Keltner, Tracy, Sauter, Cordaro, & McNeil, 2016; Parr & Waller, 2006; Waller & Micheletta, 2013). This cross-species continuity supports the idea that facial signals have evolved to serve broad social functions. The function of emotion seems to have evolved from a physiological to a communicative one, allowing others to understand one's emotional state. Both animals and humans base a lot of their decisions on cues and signals, such as the choice of their sexual partner (Thornhill & Gangestad, 1999); engaging in cooperative or altruistic behaviours (Schug, Matsumoto, Horita, Yamagishi, & Bonnet, 2010); or the appropriate social response based on the context (Blair, 2003; Shariff & Tracy, 2011). Most, if not all, fitness-related matters are linked intrinsically to signals, and thus the fitness of an individual (regardless of its species) is based on the ability to read and interpret correctly signals and cues. The survival of a being depends on the honesty of their relatives and the reliability of the signals sent. From this results a need for morality (prompting the display of honest signals) and the ability to read efficiently the signals sent by conspecifics.

#### 1.4. Morality and moral emotions

##### **Ontogenetic development of morality**

Morality can be seen as a set of rules to follow in a given society to respect the integrity of each member of the group and live in osmosis (or at least relative harmony) with each other. It is said to dignify and elevates humans above all other animals (Haidt, 2003). It is composed of moral reasoning, moral decision, moral behaviours, and judgements (Tomasello & Vaish, 2013).

The main function of morality has been argued to be to regulate social interactions between individuals and to promote cooperation within groups (Tomasello & Vaish, 2013). The concept of morality seems to appear early on in childhood (Piaget, 1965); research has shown that young children distinguish between equal and unequal distribution (M. F. Schmidt & Sommerville, 2011) and are able to discriminate between fair and unfair individuals, preferring the fair individual for future interactions (Geraci & Surian, 2011). In turn, young children were shown to be more generous to individuals that previously exhibited generous behaviours towards others (Olson & Spelke, 2008). However, those displays of early morality could be seen as based on personal relationships more than on a general understanding of moral actions (Darwall, 2006). A deeper understanding of norm-based morality has been argued to appear later in childhood, but as early as 3 years of age (Rakoczy, 2008; M. F. Schmidt & Sommerville, 2011; Wyman, Rakoczy, & Tomasello, 2009). At this age, children start to have a deeper, more general, understanding of what is right and what is wrong and recognise morality as a mutual expectation of implicit agreement regarding how one ought to behave in society (Tomasello & Vaish, 2013). 'Moral behaviours' could thus be defined as behaviours respecting morality and acting in accordance with moral standards. As imperfect human beings, behaviours are not always in line with moral standards, and we easily slip and trip along the way (Tangney, Stuewig, & Mashek, 2007).

To enforce and arouse morality since the youngest age, different modalities seems to have been used: punishments or sanctions (Turner, 2000; p.48), reputations (Haley & Fessler, 2005; Sperber & Baumard, 2012), and bedtime stories (Lee et al., 2014; Talwar, Yachison, & Leduc, 2015). What is a fairy tale if not a story meant to prompt good behaviours in children? Most of the classic tales, from the Grimm's brother to Perrault, depict the consequences of good and bad behaviours, and each of them ends with a moral.

## **Cooperation and punishment**

Humans are uniquely cooperative (Tomasello, 2008), and form crucial short- and long-term relationships between individuals (for review, see Silk, 2005). Friendships and social interactions involve cooperative interactions, often separated in time and space (reciprocal, and even delayed, altruism; Hamilton, 1964; Trivers, 1971), based on the memory of past interactions and the emotional load associated with them (Seyfarth & Cheney, 2012; Silk, 2005). The costs associated with delayed altruism are higher than those associated with direct cooperation (there is a risk that the partner will not return the favour; Mayr, 1961; Tinbergen, 1963), necessitating the simultaneous evolution of a control system. Punishment and spite have the potential to maintain cooperative behaviour and ensure equity within the relationship (Henrich & Boyd, 2001; Jensen, 2010; Jensen, Hare, Call, & Tomasello, 2006) by imposing costs on defecting partners. The costs could be the refusal to cooperate at a later time or damage to the defecting partner's reputation at the scale of the whole social group (Haley & Fessler, 2005), leading to wider repercussions against the untrustworthy. To regain the good graces of their partners, defectors need to acknowledge their wrongs, make amends (Bateson, Nettle, & Roberts, 2006), and even express remorse regarding their wrongdoing (Rosenstock & O'Connor, 2018). Moral emotions are, therefore, intimately linked to our relationships with others (Haidt, 2003), facilitating social interactions and important relationships (Keltner & Gross, 1999; Muris & Meesters, 2014).

## **Moral emotions**

Moral emotions are emotions elicited by a transgression or obedience to morality in association with moral behaviours. Emotions involved in such moral contexts are referred to as 'moral emotions' and are defined as emotions arising from a position-taking perspective, allowing one to understand how others feel and how they see the agent (Haidt, 2003). From this definition, it is clear that moral emotions are secondary emotions as to be elicited they require



a developed self, a clear distinction between self and others (Lewis et al., 1989; Tangney, 1999), as well as an understanding of standards against which the self is gauged. Since there is need of a self to be conscious of, many emotion theorists argue moral emotions would not emerge before 15 to 24 months of age (Lewis, 1995a, 1995b; Lewis, Alessandri, & Sullivan, 1992; Lewis et al., 1989), when the self begins to be defined.

In lay terms, moral emotions drive us to do good and prevent us from doing bad (Kroll & Egan, 2004; Tangney et al., 2007). They are self-conscious emotions as referring to the self, but they can be intimately linked to our relationships with others, thus also falling under the 'social' categorisation of emotion. The conscience of one's judgment over the agent's action elicits pride, shame, embarrassment, or even guilt (Haidt, 2003). Such emotions are thought to facilitate the complex journey of social interactions and relationships (Keltner & Gross, 1999; Muris & Meesters, 2014), allowing one to reflect on behaviour in light of social norms and the differing perspectives of others.

### **Shame, guilt, and embarrassment – what is the difference?**

Three emotions categorised as social/moral aroused a lot of interest through time (e.g., Keltner, 1995; Keltner & Anderson, 2000; Keltner & Buswell, 1996; Leary, Landel, & Patton, 1996; Teroni & Deonna, 2008): shame, guilt, and embarrassment. The question that arises from the study of those emotions is inevitably how to differentiate them and how do they contribute to our social life.

Embarrassment has been proposed to be the first trace of the emergence of self and morality (Tracy & Robins, 2004), whereas shame and guilt are seen as more complex self-conscious emotions and thus emerging later in development. If embarrassment emerges within the second year of life, shame and guilt are thought to surface as late as the end of the third year (Izard, Ackerman, & Schultz, 1999; Lewis, 1995a, 1995b). Embarrassment is related to

hierarchical interactions, as it is more often felt when surrounded by people of higher social status than when around people of lower social status. It is also associated with the violation of a social convention (Keltner & Buswell, 1996) and argued to be a phylogenetically older and simpler version of shame, or 'protoshame' (Fessler, 1999; Keltner, 1995).

Embarrassment is associated with the violation of social norms; what about shame and guilt? Shame and guilt have been shown to share similar attributes: they both relate to the sense of violation of social and/or moral norms (for review, see Miceli & Castelfranchi, 2018; Tangney & Tracy, 2012); strongly correlate with each other; and tend to coexist (Ferguson & Crowley, 1997). If those two emotions are confused more often than not, they appear to result from different psychological systems, with different elicitors and action tendencies (Baumeister, Stillwell, & Heatherton, 1994; Tangney, 1991). How do we then differentiate them?

The following statement can be used to illustrate the difference between these closely related emotions: "*I have done something wrong*". There are two subjects of focus: the agent (who did it) and the action (what has been done). Looking at the agent, the statement refers to the self "I", falling under the self-conscious category. The difference between shame and guilt will be made based on the answer to this question: does this action define who I am as a person? If the answer is "yes", it means the wrong you did is part of who you are, the action makes you a bad person and you tend to feel shameful about this action. This emphasises the idea that shame hurts more than embarrassment because it highlights a defect in the inner self; it appears linked to hierarchical interactions (Davidson, Scherer, & Goldsmith, 2009). If the answer is "no", the action was an isolated act that does not define you as a person and you would rather feel guilty about the wrong done. Moreover, guilt feelings tend to appear mainly in the context of communal, egalitarian, relationships (Davidson et al., 2009). It focuses on the wrongdoing done to others rather than the consequences for the self, judging one's action as bad (Lewis, 1993). Guilt seems thus to imply a negative moral self-evaluation, where someone's behaviours are

responsible for the wrongdoing and are often associated with empathy and perspective-taking (Miceli & Castelfranchi, 2018). Shame, however, seems to be associated with a negative non-moral self-evaluation, where the responsibility issues are not looked at, but rather focus on the discrepancy between one's ideal self and one's actual self (Miceli & Castelfranchi, 2018). The results of such feelings, i.e. the action tendencies resulting from feeling ashamed or guilty, differ as well. When feeling shame, the person will rather try to find a hole to hide, withdrawing from any social environment (Lewis, 1993). Such a submissive attitude indicates that the violation has been acknowledged and would thus reduce the likelihood of retaliation (Haidt 2003). In contrast, in guilt, the person will try to repair the wrong done and make amend (Haidt, 2003; Taylor, 1996).

Now looking at the action "**something**", a new question arises: do I wish I could change it? If you feel bad about what you did but you have no wish to change it, regret emerges. However, if you feel bad and you wish you could erase what you did, guilt could turn into remorse. Both guilt and remorse are action-oriented, motivating reparation by making amend or thinking you deserve to be punished (Taylor 1996) (see Figure 1.1 for an overview).



**Figure 1.1. Differentiate shame, guilt, remorse and regret.**

When the realisation of a person wrongdoing hits her, the focus is often on either the agent (“I”) or the action (“something”) and our conscience makes us reflect on whether the action define who we are (shame vs guilt) or whether we wish we could change what just happened. (remorse vs regret).

Cultural style has been defined on the relative importance of guilt and shame in the society (Benedict, 1946); Western cultures tend to be defined as guilt-culture whereas East-Asian cultures tend to be defined as shame-cultures. Moreover, Western cultures tend to have a more negative attitude towards shame, whereas Eastern societies value this emotion, and effectively induce it in others to motivate prosocial behaviour (Bedford & Hwang, 2003). Protoshame may have thus evolved differently in East-Asian and Western societies, based on the very different concepts of self and sense of community, specific to each culture (Cousins, 1989). Embarrassment/protoshame in East-Asians might be simply triggered by being in the presence of someone higher in the hierarchy and shame will be the central emotion of moral regulation (Haidt, 2003). In modern Western societies, there is a partial separation between the social order and the moral order (Turiel, 1983), resulting in a clear distinction between embarrassment, often reported when one violates social conventions, and shame, more related to violation of moral norms (Keltner and Buswell, 1996). Embarrassment is thus perceived as not

as unpleasant as shame (Haidt, 2003), uncomfortable smile or laughter are usually associated with embarrassment episodes, whereas shame hurts more deeply (Davidson et al., 2009).

Having a clearer idea of what makes shame, guilt, and embarrassment distinct emotions, what could be the diverse functions specific to guilt?

## 1.5. The case of guilt

### **What is the function of guilt?**

Guilt is one of the most social, other-oriented emotions that people can experience throughout life (J. Carroll, 1985; Cryder, Springer, & Morewedge, 2012; Tangney, 1999). Because one feels guilty over an action directed toward another person, the agent may try to make amends and repair the potentially broken relationship with this other person (Cryder et al., 2012; Tignor & Colvin, 2019). The experience of guilt is inextricably linked to the interests of another person, and the goal of maintaining, mending or assessing relationships with others. It has been shown to have a potentially positive function within social interaction, of stimulating prosocial behaviours towards and from others. For example, guilt does not prompt general behaviours to repair one's reputation but instead promotes actions towards those who have been wronged specifically (Cryder et al., 2012). Experiments have shown that guilt can prompt people to specific actions towards others, such as helping behaviours (Zemack-Rugar, Bettman, & Fitzsimons, 2007) and prejudice-reducing behaviours (Amodio, Devine, & Harmon-Jones, 2007). The intensity of the guilty feelings directly depends on the nature of the relationship with the person wronged; the closer people are to the victim, the more intense their feelings (Baumeister et al., 1994). Guilt can be experienced automatically after committing a social transgression (self-induced), but can also be induced by others as a method of control to gain

power within relationships (other-induced; Baumeister et al., 1994). Guilt can thus be a complex and powerful phenomenon within social negotiations.

However, some research has also focussed on the maladaptive function of feeling guilt and the link between the propensity to experience guilt and negative personal and interpersonal tendencies (Jones & Kugler, 1993). Any emotion, positive or negative (moral, social, or else), can be either adaptive and functional or maladaptive and dysfunctional based on frequency and appropriateness of experience. In other words, the experience of an emotion does not have a function *per se*; it is when placed in the specific context of occurrence that we can determine the function of emotion (Scherer, Mortillaro, & Mehu, 2013). Experiencing guilt in the absence of specific eliciting contexts (e.g., transgression, wrongdoing) could be perceived as inappropriate and in turn, be maladaptive. However, expressing guilt after wronging someone is socially the appropriate response and lead to sustainable relationships (Cryder et al., 2012). Moreover, Tignor and Colvin (2019) highlighted that the methodology used to measure trait guilt will impact whether it is associated with positive or negative personality-related traits (e.g. altruism, agreeableness or neuroticism, depressivity). Two types of measures can be used to measure trait guilt: situated or unsituated measures. Situated measures (e.g., Guilt and Shame Proneness Scale, GASP; Cohen, Wolf, Panter, & Insko, 2011) assess trait guilt using hypothetical scenarios of transgression and people need to report the likelihood of experiencing various emotions, including guilt. Unsituated measures (e.g., Positive and Negative Affect Scales; Watson, Clark, & Tellegen, 1988) assess trait guilt using a list of feelings and emotion words and ask people how often they experience each of them over a given period of time (e.g. over the past week, over the past month). Unsituated measures of guilt have been found to correlate with maladaptive social traits (e.g., neuroticism, depressivity) whereas situated measures were correlated with adaptive personal traits (e.g., altruism, agreeableness; Tignor & Colvin, 2019). However, unsituated measures have also been used as mood check (PANAS), assessing state emotions, i.e. how is one feeling right now in the present situation. It seems thus that when

used as such, unsituated measure of state guilt could be correlated with adaptive personal and social traits.

### **Facial signals of guilt**

Facial expressions could have evolved in a cooperative context not only as a display of emotion (Ekman & Friesen, 1971) but also as a prediction of future behaviours (Fridlund, 1994). The BET (Ekman, 1971) argues for facial expression to be signals indicating the emotions or internal state of the sender. As such, displaying guilt might not be adaptive, as it would show weakness and acknowledgement of a wrongdoing from the sender. The BECV (Fridlund, 1994, 2017) argues for facial expressions to be signals sent by the sender to indicate to the receivers his/her most likely future behaviours. Facial expressions are not necessarily the readouts of emotions but part of plans of action in social interactions. As such, a facial expression of guilt could have evolved in humans due to the value in indicating one's willingness to make amend. Being able to detect and understand such guilt display would be valuable for the receiver, to acknowledge the good intention of the wrongdoer. It is then possible that guilt can be communicated via the face as a universal facial expression, which could function as an honest signal (K. L. Schmidt & Cohn, 2001) sent by the wrong-doer indicating his/her willingness to appease the receiver and repair the damaged relationship (Fridlund, 2017), serving an appeasement function (Keltner & Anderson, 2000).

As secondary emotion, guilt can be heavily influenced by the socio-cultural context, and previous research found marked differences between cultures (Bedford, 2004; Bedford & Hwang, 2003). However, it is not known whether these differences are restricted to feelings/experiences of guilt, or also relate to its behaviour and function within social interactions. Most cross-cultural research in production (Matsumoto, Yoo, & Fontaine, 2008) and perception (Elfenbein et al., 2007; Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Jack,

Caldara, et al., 2012; Jack, Garrod, et al., 2012) of facial expressions of emotions compare Western and Asian cultures. Individuals in East Asian cultures have been shown to be more sensitive to their close relatives' guilt than individuals in European cultures (Cousins, 1989), and also to exhibit greater motivation to minimise disruptions to relationship harmony (Kim, Sherman, & Taylor, 2008).

However, it is not known whether the two cultures differ only in degree (how much), or whether they are also differing in their proneness to self-induced vs other-induced guilt. Chinese cultures are often structured by a sense of duty, where individuals are perceived as part of something bigger (family, company, country); when Western cultures tend to be based on a sense of self, where individuals are distinct entities praised for their uniqueness. Considering those differences, it could be hypothesised that individuals that grew up in an Asian society would be more prone to other-induced guilt than an individual raised in a Western society, who would then be more sensitive to self-induced guilt. Moreover, based on linguistic data different kinds of guilt were distinguished in Asian cultures (Bedford and Hwang, 2003), some that are not even defined in Western cultures, usually referring to the reason to feel guilty. Finally, as Chinese culture is based on restraint of expression and self-control, they might also exhibit less facial expressions of guilt than Western cultures, despite reporting stronger feelings of guilt and greater motivation to minimise disruptions (Kim et al, 2008).

To date, there is no evidence that an expression of guilt is a genuine behavioural phenomenon. We do not currently understand the mechanism and function underlying the common impression that we can tell when someone feels guilty. We have no data on whether typical and specific non-verbal behaviours accompany feelings of guilt, whether others can see such expressions as reliable indicators of guilt, and whether this process affects social outcomes. We do not know either if different cultures rely on similar cues to detect guilt on a face; there may be a recognisable facial expression shared by all cultures that also exhibit cultural and



individual variation. Yet to date no facial movement has been reliably associated with the experience of guilt.

This PhD aimed to explore the emotion of guilt from three intertwined perspectives: the behavioural expression of guilt, the experience of guilt, and the responses to guilt from observers. More specifically, I aimed to identify whether there are reliable facial traits associated with reported feelings of guilt that could act as a signal detectable by others. I was also interested in looking at the consequences of guilt: what do you do when you feel guilty and how do you react when spotting guilt on someone's face? I tried to provide some answers to those questions throughout this project. I also aimed to understand better the true profile of human guilt by adopting a cross-cultural perspective, comparing Chinese populations to British-born individuals.

The research questions I will address through this PhD are:

1. Is there a facial signal of guilt?
  - a. Are there facial movements reliably associated with self-reported feelings of guilt?
  - b. Do individuals from different cultures express guilt using similar movements?
2. Is there a facial signal reliably identified as guilt by social partners?
  - a. Do facial signals produced during the experience of guilt have a communicative message understood as guilt by a social partner?
  - b. Is contextual information necessary to be able to spot guilt in people displaying those patterns of movements?
3. Is the experience of guilt modulated by the quality of social interactions?
  - a. Does the intensity of the feelings/expressions depend on who you feel guilt towards (friend or stranger)?
  - b. Are your actions the same in different situations?
    - i. Self- vs other-induced?
    - ii. Friend vs stranger?

## 2. Are there non-verbal signals of guilt?<sup>2</sup>

### Abstract

Guilt is a complex emotion with a potentially important social function of stimulating cooperative behaviours towards and from others, but whether guilt is associated with a pattern of nonverbal behaviour recognisable as guilt is unknown. We examined the production and perception of guilt in two different studies, with a total of 238 participants with European and East-Asian backgrounds. Guilt was induced experimentally, eliciting patterns of movement that were associated with both the participants' subjective feelings of guilt and judges' impressions of their guilt. Self-reported guilt was most closely associated with upper lip movement and neck touching, but the perception of guilt by others was most closely associated with frowning and face touching. While there were differences in production and perception, both are characterised by self-directed behaviours. Cultural differences were minimal but included some differences in facial movements of the eyes. The findings suggest that there is a basic non-verbal pattern of guilt in humans.

**Keywords:** Facial expression, guilt, emotion, FACS, culture, self-directed behaviour

---

<sup>2</sup> Results from those studies have been presented at the Consortium of European Research on Emotion (Leiden 2016 and Glasgow 2018) and at the International Society for Research on Emotion (Saint-Louis 2017). A version of the chapter has also been submitted to PLOS ONE and is currently under review.

## 2.1. Background<sup>3</sup>

Humans are highly cooperative with both relatives and strangers (Tomasello, 2008), and the need for cooperation may have provided a powerful selection pressure behind many of the behaviours that we consider uniquely human. Guilt is an emotional and cognitive experience arising when someone feels that they did something wrong. It could have evolved due to its potential adaptive function, within social interaction, of promoting social tendencies towards and from others. Guilt is classified as a moral, self-conscious emotion, along with pride, shame and embarrassment (Haidt, 2003), and is one of the most social, other-oriented emotions that people experience throughout life (J. Carroll, 1985; Cryder et al., 2012). Despite a growing literature on the social consequences of feeling guilty (Cryder et al., 2012; De Hooge, Zeelenberg, & Breugelmans, 2007; Ketelaar & Au, 2003), little is known about the behavioural mechanisms underpinning the social functions of guilt, specifically whether guilt is associated with a specific facial expression or nonverbal signal that others can recognise. If people can recognise guilt in others, this might explain how guilt can facilitate cooperation and pro-social behaviours within social interaction. People often state that they can detect a feeling of guilt in others (Weisman, 2014), but with some notable exceptions (Keltner & Buswell, 1996), to date, a facial expression of guilt has not been identified scientifically.

Moral emotions are thought to facilitate the complex navigation of social interactions and relationships (Keltner & Gross, 1999; Muris & Meesters, 2014; Vaish, 2018), allowing one to consider behaviour in light of social norms and the differing perspective of others. Early accounts of guilt cast it primarily as a self-regulatory emotion prompting individuals to reflect on their mistakes and ultimately feel better about themselves (Cryder et al., 2012). It has been shown to have a potentially positive function within social interaction of triggering pro-social tendencies

---

<sup>3</sup> The literature review covered in the empirical chapters might overlap between chapters, as well as with the general introduction, as each chapter is being prepared for submission in peer-reviewed journals.

towards and from others, promoting actions towards those who have been wronged specifically (Cryder et al., 2012; De Hooge et al., 2007; Ketelaar & Au, 2003). Experiments have shown that guilt can prompt people to specific actions towards others, such as helping behaviours (Zemack-Rugar et al., 2007) and prejudice-reducing behaviours (Amodio et al., 2007). The virtue of apologies and verbal admittance of wrong-doing has been well studied (O'Malley & Greenberg, 1983; Ohbuchi, Kameda, & Agarie, 1989), revealing that people (victims of wrong-doing or bystanders) would be more lenient towards a wrong-doer that recognise their faults. For example, in legal contexts, judges and jurors claim that they know when a defendant is sorry for the crime they have committed (Weisman, 2014), which can then impact on sentencing. Guilt, however, is not just a social emotion. Indeed, a *Dobby Effect* has been highlighted (Nelissen & Zeelenberg, 2009), refuting the all-social aspect of guilt, and showing that guilty people sometimes punish themselves in the absence of opportunity to make amends to the victim of their wrong-doing. The social aspect of guilt seems then linked to the context the guilty person finds themselves in: they will act pro-socially and make amends in social context but will engage in self-punishment when socially isolated (Nelissen & Zeelenberg, 2009). Finally, guilt can be experienced automatically after committing a social transgression (self-induced), but can also be induced by others as a method of control to gain power within relationships (other induced; Baumeister et al., 1994). Guilt can thus be a complex and powerful phenomenon within social negotiations, but whether guilt can be observed by others without being explicitly declared is unknown. If guilt can be detected in this way, the potential to affect social outcomes between individuals is increased.

Whether emotions (and which emotions) are associated with universally produced and recognised facial expressions is debated. The classic and largely dominant view, the Basic Emotion Theory (BET; Crivelli & Fridlund, 2019; Ekman & Cordaro, 2011; Ekman & Friesen, 1969), is that primary, basic emotions (happiness, sadness, anger, surprise, disgust and fear - Ekman & Cordaro, 2011) are considered innate to all human populations and universally expressed

(Brown, 1991; Ekman & Friesen, 1969), and so likely resulting from specific functional adaptations (Ekman & Keltner, 1970). In contrast, secondary emotions (of which guilt is one, along with embarrassment, shame, and contempt) are thought to differ significantly between cultures (Bedford & Hwang, 2003; Matsumoto et al., 2005), their expressions subject to specific cultural display rules (Ekman & Friesen, 1969; Matsumoto et al., 2005), and acquired and developed gradually during childhood (Tangney, 1999). Secondary emotions are more idiosyncratic and context-dependent, which is why it has been difficult to identify specific facial movements associated with the experience of those emotional states. This later ontogeny led scientists to explore the possible influence of environment on the development of secondary emotions (Ekman & Friesen, 1969) and through the impact of these variable environments they are not thought to have a prototypical universal expression (Izard, 1994). Within a Behavioural Ecological View of facial expressions (BECV; Fridlund, 1994, 2017), however, the distinction between primary and secondary emotions is less rigid. BEV argues that facial expressions indicate the sender's most likely future behaviours and thus function as important social signals in social interaction. Facial expressions benefit both the sender and receiver by reducing the need for conflict when interests are declared openly (Baumeister et al., 1994; Ekman & Friesen, 1971; Fridlund, 1994, 2017; Tomasello, 2008; Waller, Whitehouse, & Micheletta, 2016). As such, both primary and secondary emotions *can* be associated with specific, readable and recognisable, facial signals, as it is not the emotion per se that is being transmitted, but instead the potential social action (Fridlund, 2017; Waller et al., 2016). Therefore, if guilt is associated with a specific social outcome (e.g. making amends), it is possible that others can detect this from nonverbal behaviour, specifically from a facial signal with communicative value. Signals can, therefore, be understood as a way for an individual to manipulate or alter the behaviour of another individual (Krebs & Davies, 1993; Krebs & Dawkins, 1978). Signals can also be used by others when deciding if and how to respond to a given situation (Bradbury & Vehrencamp, 1998). The potentially important role of the face in social interactions led us to hypothesise that guilt would be associated with an identifiable facial, or at least non-verbal, signal.

A non-verbal signal could include not only facial expressions (i.e., resulting from the contraction of specific facial muscles), but also head position and behaviours directed towards the head (e.g., touching the face or hair), body postures, or gestures. Non-verbal behaviours (facial expressions and actions directed towards the face) would be considered a signal if those behaviours are reliably associated with the experience of guilt and are accurately perceived by observers as an indication of guilt, as well as influencing the observers' behaviours (Dezecache et al., 2013). We tried here to identify non-verbal signals resulting from a specific cognitive appraisal (i.e., a situation designed to induce guilt; Scherer et al., 2013), occurring concomitantly with a self-reported feeling of guilt. By doing this, we are following Scherer et al. (2013)'s view that non-verbal signals can carry emotional meaning, as well as action tendencies which can both be perceived and interpreted by observers. Moreover, some researchers argue that the concept of emotion is constructed (Barrett, 2017a, 2017b; Barrett, Mesquita, & Gendron, 2011), as the result of a given experience, at a specific time, in a specific context (Barrett, 2017a). As such, both theories (Barrett, 2017a; Scherer et al., 2013) advocate for a less direct link between non-verbal signals and emotional states than previously argued by the Basic Emotion Theory (Ekman & Cordaro, 2011).

Some secondary emotions (e.g. shame and embarrassment; Ekman, 1971; Keltner & Buswell, 1996) have been associated with recognisable facial movements, but these emotions are often confused with each other. Guilt can also be mistaken or mislabelled as shame, and sometimes embarrassment and research has tried to differentiate between those, not only in terms of the psychological meaning but also in terms of the behavioural signal (Keltner & Buswell, 1996; Tangney, 1999; Teroni & Deonna, 2008). The specific social context in which the facial expression is placed therefore can be important in interpretation of these expressions (Aviezer, Hassin, Bentin, & Trope, 2008; Hess, Blaison, & Kafetsios, 2016). Nevertheless, there

must be some key physical elements to such expressions that underpin their recognition in order to make them in some way identifiable to others.

For instance, the action tendencies of shame, embarrassment and guilt, are rather different, and may thus manifest as physical differences in a behavioural signal. Behavioural responses to embarrassment and shame have been clearly identified over the years (Izard, 1977; Keltner, 1995; Lewis et al., 1992): embarrassment displays are marked by gaze down, controlled smiles, gaze shifts, and face touches (Keltner, 1995) whereas a shameful display is marked with head and gaze down (Izard, 1977; Keltner, 1995; Lewis et al., 1992). Embarrassment serves a reconciliatory and appeasement function, reconciling in social relations following transgressions (see Keltner & Anderson, 2000 for review) whereas shame serves a reconciliatory and appeasement function following hierarchical transgressions. In contrast, a facial expression of guilt has not been clearly described. Although guilt may have evolved in humans due to the value in indicating one's willingness to make amends. Only one study has tried to identify a recognisable set of facial movements associated with the experience of guilt (Keltner & Buswell, 1996). Three potential displays of guilt were presented on still photographs: a facial expression representing self-contempt, which has been shown to be associated with the experience of guilt (Higgins, 1987); a non-verbal display of sympathy (Eisenberg et al., 1989), which could be part of the experience of guilt; and finally a facial expression of pain, considered as one antecedent of guilt (Emde, Johnson, & Easterbrooks, 1987). Following the presentation of the still photographs, participants in this study had to select one emotion word among 14 different options (including a "no emotion" option). None of these conceptualised displays of guilt was identified as such by observers (Keltner & Buswell, 1996). The authors speculated that participants may have struggled with identifying fixed displays compared to spontaneous dynamic stimuli of the same emotions (Keltner, 1995; Keltner & Buswell, 1996). This study motivated us to try a new methodology, with a bottom-up approach to try inducing guilt in the



laboratory in order to collect spontaneous dynamic displays associated with the experience of guilt that we could then present to naïve observers.

Individuals in East-Asian cultures are more sensitive to their close relatives' guilt than individuals in European cultures (Cousins, 1989), and also exhibit a greater motivation to minimise disruptions to relationship harmony (Kim, Sherman, & Taylor, 2008). Cultural differences (or accents; Marsh et al., 2003) would not necessarily contradict with an evolutionary view. Different cultures may rely more or less on specific facial cues, which could be functional in a given culture and for a given emotion, but not necessarily in other cultures (Jack, Caldara, et al., 2012; Keller & Otto, 2009). Moreover, a given emotion could be more common in one culture (i.e., anger in European countries - Keller & Otto, 2009) or people in different cultures might experience emotions differently (i.e., anger associated with a feeling of control in European countries but a feeling of guilt in East-Asian countries - Keller & Otto, 2009). Previous research also demonstrated that individuals from different cultures used different cues to identify emotional facial expressions (Jack, Caldara, et al., 2012; Yan et al., 2017) or produce different facial expressions in the same situation (De Leersnyder & Mesquita, 2015). However, it has also been shown that time spent in a culture improved the recognition of a culturally specific facial expression (Elfenbein & Ambady, 2003), suggesting that people can become acculturated to different styles of emotional recognition. Additionally, recent work suggests that if people change cultural environment they can adapt to their new culture by experiencing emotion according to their new cultural standards (De Leersnyder, 2017). The cultural impact on facial expressivity and perception is thus not limited to early childhood exposure but can influence adults during their lifetimes.

### **Present Investigation**

Here, we examined variation in both the production and perception of the specific facial movements associated with guilt in a culturally varied, recruiting people from WEIRD and non-

WEIRD countries (Western, Educated, Industrialised, Rich and Democratic societies; Henrich, Heine, & Norenzayan, 2010). As individuals in East-Asian cultures are more sensitive to their close relatives' guilt than individuals in European cultures (Cousins, 1989), there may be a recognisable facial expression shared by both cultures but that also exhibits cultural and individual variation. We examined the production and perception of spontaneous facial expressions using a bottom-up approach to identify dynamic patterns in facial behaviour, departing from the classic method of coding the apex of expression or movements of interest only (Galati et al., 2003; K. L. Schmidt, Cohn, & Tian, 2003). We looked at the production of facial movements in individuals belonging to two different cultures in order to assess whether patterns produced could be considered cross-cultural or were tied to one given culture in this particular context. Firstly, we identified facial movements based on what people displayed when experiencing guilt. Secondly, we identified facial movements based on what people perceived as guilt. This study looked at the production of a facial expression of guilt using for the first time an experimental induction approach and an extensive dynamic facial movement coding system.

## 2.2. Study 1 – Production of guilt:

### Methods

#### Participants

One hundred and thirty-one participants took part in this study (94 females;  $M_{age} = 25.41$ ,  $SD = 9.47$ ; see Table 2.1 for details). Participants were recruited based on an opportunistic sampling method and were all UK resident at the time of the experiment (but included both UK and international individuals). All of them received either course credit (if student) or £5 for their time. The whole experiment lasted 45 minutes on average. Our sample consisted of two distinct populations. Population 1 (72 participants who consisted of predominantly European University students and staff), and Population 2 (59 participants

who consisted of predominantly overseas students of Asian origin; for a breakdown of the whole sample, Table 2.1). Belonging to Population 1 or 2 resulted from self-identification as participants reported the cultural background they identified to. Participants within the European population had stayed in the UK ( $M = 25.05$  years,  $SD = 14.06$ ) for significantly longer than participants within the East-Asian population ( $M = 2.88$  years,  $SD = 4.98$ ;  $t(125) = 10.716$ ,  $p < 0.001$ ). Most East Asian participants had recently arrived in the UK to study at the University (see Table 2.1 for details). The project has been reviewed and approved by the Science Faculty Ethics Committee (SFEC) from the University of Portsmouth (Appendix 2). Each participant signed an informed consent form granting authorisation for the use of the data for research purposes. The individuals in this manuscript (Figure 2.2) have given written informed consent (as outlined in PLOS consent form) to publish these case details.

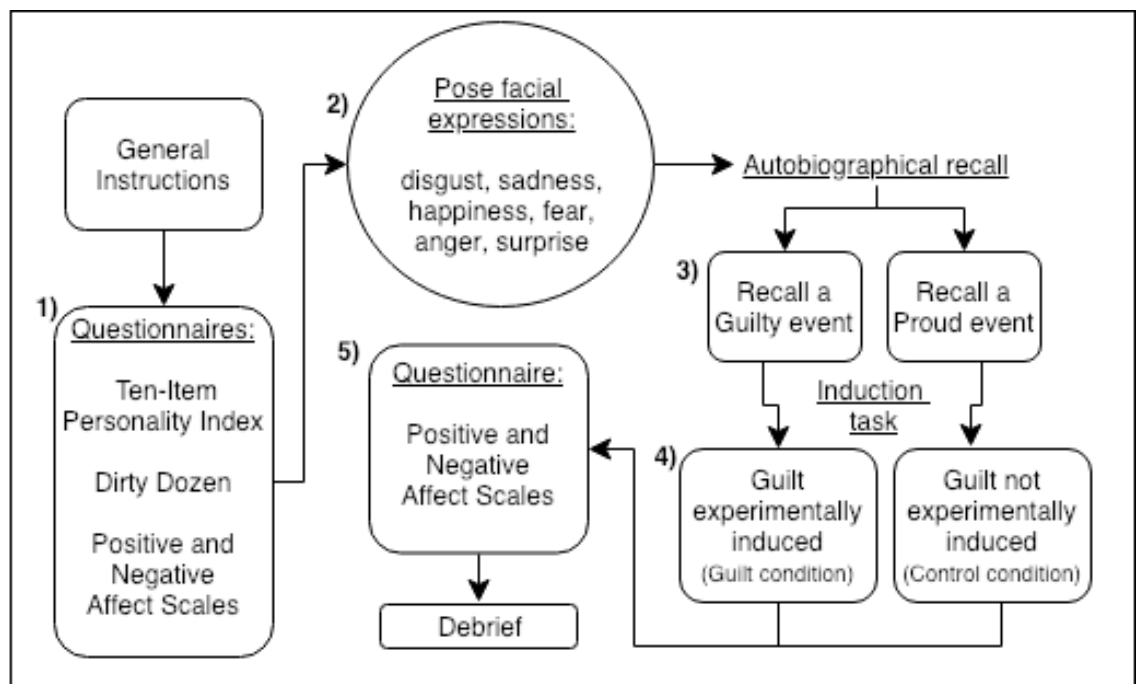
**Table 2.1. Participants demographic information (Study 1)**

<b>Nationality</b>	Count	Average time spent in UK
<b>European</b>	<b>71</b>	<b>24.55</b>
British	63	26.58
British & German	1	20.00
Czech	1	0.60
Dutch	1	12.00
Estonian	1	17.00
French	2	1.75
Italian	1	21.00
Nigerian	1	0.20
<b>Asian</b>	<b>57</b>	<b>2.89</b>
British	4	16.25
British & Chinese	1	16.00
Chinese	25	1.22
Filipino	1	17.00
Hong-Kong	5	0.88
Hong-Kong (British National Oversea)	1	2.00
Indian	1	4.00
Indonesian	3	1.69
Japan	1	2.00
Malaysian	6	0.55
Singaporean	2	0.55
South Korean	1	0.50
Tibetan	1	2.00
Vietnamese	5	2.20
<i>Grand Total</i>	128	14.77

*N.B. Time spent in the UK presented in years; participants self-identified as "European" or "Asian" and their nationalities are presented in this table.*

### General Procedure

To begin, participants were given general instructions regarding the experiment and written consent was obtained. Participants were originally told that this study had a different aim - to assess how personality affects behaviour and facial expressions. Following these instructions, the rest of the tasks were displayed on a computer using the OpenSesame© software (Mathôt, Schreij, & Theeuwes, 2012), and the participant was filmed for the remaining time (using a JVC Everio GZ-MG750, 25 frames/second, placed approximately 50 cm away from their face). The experiment consisted of 5 key steps, as outlined in Figure 2.1 and explained in more detail below.



**Figure 2.1. General Procedure.**

A flowchart representing the procedure of the experiment in Study 1.

First (Figure 2.1: 1), participants were required to complete two personality questionnaires, the Ten-Item Personality Index (TIPI; Gosling, Rentfrow, & Swann, 2003) and the Dirty Dozen (DD; Jonason & Webster, 2010), followed by a mood-check questionnaire (Positive and Negative Affect Scales - PANAS; Watson et al., 1988. For details on the questionnaires, see Appendix 1). Question order was randomised between individuals. Personality questionnaires were used as a validation of our cover-up story (examining the impact of personality on behaviours and facial expressions), and to investigate whether main personality traits correlated with self-reported guilt. Second (Figure 2.1: 2), participants were prompted to pose six emotional facial expressions (anger, fear, disgust, happiness, surprise, and sadness) in succession and hold each of them for 8 seconds. This acted as a distractor task to disassociate the participants from the previous questionnaires. We used a video camera to record facial expressions of emotional states produced in this task. The experimenter then pretended to turn off the camera, but, in reality, kept recording the rest of the experiment (note that they were fully debriefed at the end of the experiment). Third (Figure 2.1: 3), participants were asked to recall an autobiographical event, and write about this in some detail. Participants were prompted to either recall an event where they felt guilty regarding something they did for relative (guilt condition; N = 66, 28 East-Asians and 38 Europeans), or an event where they felt proud (control condition; N = 65, 31 East-Asians and 34 Europeans). Participants were randomly assigned to a single condition (guilt or control). This third task was used as both a priming task to start inducing either guilt or pride in participants (as used in previous research; Rebeca, Apostol, Benga, & Miclea, 2013), and was a necessary component of the following induction task. Fourth (Figure 2.1: 4), in participants who were assigned to the guilt condition, and recalled a guilty autobiographical event, guilt was induced further experimentally. Here, the experimenter asked the participant to save their written recall on a USB flash drive. Once returned to the experimenter, the participant was informed that the USB flash-drive had become corrupt, and their data, among all the other data of other participants in the study, had been lost. During this social interaction

between the experimenter and participant, it was clearly implied that it was the participant's fault. They were told that this would be reported to the Principal Investigator and that there was nothing to be done at the present time. They were finally invited to resume the experiment. In the control condition, guilt was not induced and participants were told that their written recall was saved correctly on the USB flash-drive and were asked to continue. Finally, (Figure 2.1: 5), participants completed a second PANAS questionnaire, after which they were informed about the true aim of the experiment, told that the camera had kept on recording and debriefed.

As mentioned, previous research used autobiographical recall as an induction task, relying on the fact that remembering a previous emotional state can elicit said emotion again (Rebega et al., 2013). Using this methodology, group differences have been found between guilt-recall and control-recall (De Hooge, Nelissen, Breugelmans, & Zeelenberg, 2011; Ketelaar & Au, 2003). However, inducing an emotional state in the present is more ecologically valid and might standardise the feelings of guilt across participants to a greater extent (Rebega et al., 2013). Building on previous research (De Hooge et al., 2007), we therefore decided to use the autobiographical recall as a priming task, to get participants to start thinking about one of two emotions (pride or guilt), and then the experimental induction as a standardised induction of state guilt (or pride). We chose pride for the control condition as both pride and guilt are categorised as secondary emotions (Haidt, 2003), and we wanted to elicit in participants the same level of emotional involvement in the task in both conditions. We chose a positive secondary emotion for the control condition in order to make participants think about the recalled event in both conditions. Asking participants to recall their breakfast did not seem as strenuous or emotional as recalling a time they felt guilt/proud for something they did. The analysis of the autobiographical recalls are not presented here, but their length was similar in the guilt ( $M = 147.4$  words;  $SD = 99.36$ ) and

pride ( $M = 136.5$  words;  $SD = 79.54$ ) conditions ( $p = 0.5$ ). This made us confident that the involvement and emotional component were similar between the conditions.

### **FACS coding of facial movements**

Videos collected during the experiment were first cropped to extract the relevant moment only: the induction task (Figure 2.1, 4). This was coded for facial movements using the Facial Action Coding System (FACS; Ekman & Friesen, 1978; Ekman et al., 2002). Self-directed behaviours (face and neck touch) were also coded due to their direct links to negative affective states (Troisi, 2002). All facial movements produced by the participant during the induction task (Figure 2.1, 4) were coded for Action Units (single muscle movements; AUs) or Action Descriptors (one or more unspecified muscle movements; ADs), in both the guilt (average duration = 73.66 sec;  $SD = 46.56$  sec) and the control (average duration = 7.90 sec;  $SD = 4.27$  sec) conditions. Videos of posed emotional facial expressions were coded and used for the reliability but they were not analysed further. A full list of coded movements, defined by the FACS, can be found in Table 2.2. To simplify the coding, we did not use a 5-point intensity scale as defined in the FACS manual (Ekman & Friesen, 1978) but a dichotomous system: small (A/B) vs high (C/D/E) intensities were coded. Blushing was originally part of the ethogram but as no occurrence was observed, it was removed from further analysis. As research assistants conducted the experiment, I was able to be the main FACS coder as I was blind to the conditions and could code the videos unbiased. Moreover, we conducted reliability between the main coder (myself) and a research assistant; both coders are trained FACS coders and successfully passed the official FACS test (see below for details on reliability).

For every participant, we obtained the total frequency of different AU/ADs produced (i.e., the number of instances, from start to end, for each given AU/AD) in a given condition

as well as the overall duration (i.e., total time the AU/AD was expressed on a face) an AU/AD was produced for (Martin & Bateson, 1993). We were also able to extract temporal data, giving us the state of each AU/AD at a given frame in the video (absent, present at a small intensity or present at high intensity; 25 frames per second). All coding was conducted using the Interact© software (Mangold, 1998).

**Table 2.2. List of Action Units (AUs), movements, and miscellaneous actions coded for**

AU	AU Description	Group	Used in Study 1 PCA		Used in Study 2 PCA	
			Guilt	Control	Guilt	Control
1	Inner Brow Raiser	Upper Face AU				
2	Outer Brow Raiser	Upper Face AU				
4	Brow Lowerer	Upper Face AU				
5	Upper Lid Raiser	Upper Face AU				
6	Cheek Raiser	Upper Face AU				
7	Lids Tight	Upper Face AU				
9	Nose Wrinkle	Lower Face AU				
10	Upper Lip Raiser	Lower Face AU				
11	Nasolabial Furrow Deepener	Lower Face AU				
12	Lip Corner Puller	Lower Face AU				
14	Dimpler	Lower Face AU				
15	Lip Corner Depressor	Lower Face AU				
16	Lower Lip Depressor	Lower Face AU				
17	Chin Raiser	Lower Face AU				
18	Lip Pucker	Lower Face AU				
20	Lip Strech	Lower Face AU				
22	Lip Funneler	Lower Face AU				
23	Lip Tightener	Lower Face AU				
24	Lip Presser	Lower Face AU				
28	Lips Suck	Lower Face AU				
25	Lips Part	Lip Parting				
26	Jaw Drop	Jaw Opening				
27	Mouth Strech	Jaw Opening				
51	Turn Left	Head Position				
52	Turn Right	Head Position				
53	Head Up	Head Position				
54	Head Down	Head Position				
55	Tilt Left	Head Position				
56	Tilt Right	Head Position				
57	Head Forward	Head Position				
58	Head Backward	Head Position				
M59	Nodding	Head Position				
61	Eyes Left	Eye Position				
62	Eyes Right	Eye Position				
63	Eyes Up	Eye Position				
64	Eyes Down	Eye Position				
37	Lip Wipe	Miscellaneous				
50	Talking	Miscellaneous				
Sorry	Saying sorry	Miscellaneous				
Laugh	Laughing	Miscellaneous				
Face Touch	Touching any part of face with hand	Miscellaneous				
Neck Touch	Touching neck with hand	Miscellaneous				

\*AUs and movements used for original PCA in Study 1

#AUs and movements used for original PCA in Study 2

Note: AU10 and AU11, as well as AU23 and AU24 and AU26 and AU27, were combined for analysis



Any AU/AD that was significantly less common than the average was excluded. We used a binomial exact test as criteria for exclusion - if any AU/AD was produced by fewer participants than the calculated criteria (here, the criteria given by the binomial test was a minimum of 39 participants in both conditions), this AU/AD was not explored further in attempt to maintain a robust dataset. The binomial exact test allowed us to keep facial movements produced significantly more than chance: if at least 39 participants produced the movement, then this movement reliably occurs across participants and is not resulting from individual differences. We also ran a correlation between all AU/ADs and the self-reported guilt to select AUs that were produced more when reporting a higher level of guilt. This process (binomial test and correlations) resulted in 15AU/ADs in the guilt condition (respectively 10 in the control condition; out of a possible 39 observed in our data, see Table 2.2) and 117,781 frames in the guilt condition (12,472 frames in the control condition) for further analysis.

Reliability analysis on these 15 AUs was conducted on 5% of the video clips extracted from the videos collected during the study (42 of 820 videos, half of which were from the posed facial expression task and half from spontaneous facial expressions during the induction task, from both control and guilt conditions). Reliability analysis is important for FACS coding to ensure that the coding is unbiased and all the produced movements were observed and reported by the main coder (myself). We used the standardised measure of reliability for FACS (Wexler's Agreement), as described in Ekman, Friesen, and Hager's Investigator's Guide (Ekman et al., 2002), and agreement between coders was sufficient (Wexler's agreement = 0.708). As Krippendorff's alpha (Krippendorff, 1970) has been recommended as the "gold standard" for assessing the reliability (Hayes & Krippendorff, 2007), we calculated the Krippendorff's alpha using the "KAlpha" macro for use with IBM SPSS version 24 (IBM, 2016). Krippendorff's alpha coefficient is considered reliable if the 95% confidence was greater than chance (i.e., if the lower bound was >0). According to this index,

the reliability coefficient was significantly greater than chance ( $\alpha = 0.740$ ; K- $\alpha$  95% LCI: 0.684; K- $\alpha$  95% UCI: 0.788).

### Statistical Analysis

First, we looked at potential differences between populations in personality data. Personality indexes were compiled from the TIPI and the DD following the instructions presented in Appendix 1. Then, to assess for successful induction of guilt during the guilt induction task, we compared the data collected through the PANAS questionnaires (before vs. after induction) using a within-subjects *t*-test. We tested for a change in positive and negative affect before vs. after induction, and additionally, some specific emotional changes in guilt, shame, distress and pride (see Appendix 1 for details).

Secondly, to identify AU/ADs of interest, we ran a factor analysis (Principal Component Analysis with Varimax rotation, PCA) on the frequency data of the 117,781 frames in the guilt condition (15 AU/ADs selected) and 12,472 frames in the control condition (10 AU/ADs selected). PCA was conducted on a frame-by-frame dataset to produce factors consisting of temporally clustered AU/ADs, as previously done by Stratou, Van Der Schalk, Hoegen, and Gratch (2017). As previously shown, conducting a frame-by-frame PCA on FACS data is the optimal method to identify any link between cognitive processes and resulting factors of facial movements (Stratou et al., 2017). We added the self-reported guilt to both PCAs (guilt and control) in order to identify movements that were produced more often when guilt was reported. PCA was conducted in SPSS version 24 (IBM, 2016).

Third, we investigated the relationship between the AUs clustering with self-reported guilt (in the guilt condition only) and self-reported shame or pride using a generalized linear model analysis approach (GLM). The self-reported emotion of an individual was set as our

dependent variable, and the frequencies of AUs identified in the PCA were set as our predictor variable. We fitted the GLM using the function `glm` provided by the packages `lme4` and `lmerTest` for RStudio (Bates, Mächler, Bolker, & Walker, 2014).

Finally, we looked at the frequencies of the AUs identified by the PCAs to assess differences between the guilt and control conditions. We controlled the frequencies of production of each participant by the total number of movements produced in each video and compared the production level in both conditions.

As our sample was made up of two distinct populations (Population 1 and 2), we used this as an opportunity to assess population differences in the production of guilt. To do this we dividing our dataset into the two perceived populations, those predominantly of European origin (Population 1) and those predominantly of Asian origin (Population 2), and we ran two PCAs using the same 15 AUs/ADs previously selected and the self-reported guilt.

## Results

### Personality data

Overall, participants from different cultural groups did not differ in agreeableness, emotional stability, extraversion, openness, Machiavellianism, psychopathy, or narcissism ( $p > 0.05$ ). Participants from Population 1 presented higher levels of conscientiousness ( $M=4.199$ ,  $SD=1.213$ ) than participants from Population 2 ( $M= 3.558$ ,  $SD=1.713$ );  $t(126)=-2.462$ ,  $p=0.015$ . None of the personality traits correlated with the expressivity ( $p > 0.05$ ), which means that any difference between groups is due to a cultural impact only. Machiavellianism was positively correlated with the self-reported guilt ( $\beta = 0.0674$ ;  $SE =$

0.0241;  $p < 0.001$ ) and narcissism was negatively correlated with self-reported guilt ( $\beta = -0.185$ ;  $SE = 0.0907$ ;  $p = 0.0439$ ).

### **Guilt induction**

In our guilt condition, participants reported more negative affect after ( $M = 21.89$ ,  $SD = 8.23$ ) the guilt induction task (Figure 2.1:4) compared to before ( $M = 18.61$ ,  $SD = 8.56$ ;  $t(65) = -2.68$ ,  $p < 0.001$ ,  $d = 0.36$ ; see Table 2.3). They also experienced a decrease in positive affect after the induction ( $M = 20.27$ ,  $SD = 8.08$ ) compared to before ( $M = 29.73$ ,  $SD = 6.12$ ;  $t(65) = 9.02$ ,  $p < 0.001$ ,  $d = -0.99$ ).

More specifically, we found an increase in guilty feelings after the induction task ( $M = 2.7$ ,  $SD = 1.23$ ) compared to before ( $M = 1.35$ ,  $SD = 0.79$ ;  $t(65) = -8.31$ ,  $p < 0.001$ ,  $d = 1.28$ ). Participants also reported higher levels of shame after the induction ( $M = 2.24$ ,  $SD = 1.12$ ) compared to before the interaction with the experimenter ( $M = 1.33$ ,  $SD = 0.83$ ;  $t(65) = -5.91$ ,  $p < 0.001$ ,  $d = 1.00$ ). Participants reported a significantly higher level of guilt than shame after the induction task ( $t(65) = -3.00$ ,  $p = 0.0038$ ,  $d = 0.47$ ; see Table 2.3).

Finally, participants reported an increase in distress after ( $M = 2.42$ ,  $SD = 1.15$ ) the induction task compared to before ( $M = 1.58$ ,  $SD = 0.95$ ;  $t(65) = -5.29$ ,  $p < 0.001$ ,  $d = 0.79$ ), as well as a decreased pride after ( $M = 1.89$ ,  $SD = 1.89$ ) the induction compared to before ( $M = 2.35$ ,  $SD = 1.06$ ;  $t(65) = 3.84$ ,  $p < 0.001$ ,  $d = -0.38$ ; see Table 2.3).

Table 2.3. Self-reported emotions before and after guilt induction

Conditions	Time	Positive Affect		Negative Affect		Guilt		Shame		Distress		Pride	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
GUILT	before induction	29.73	6.12	18.61	8.56	1.35	0.79	1.33	0.83	1.58	0.95	2.35	1.06
	after induction	20.27	8.08	21.89	8.23	2.7	1.25	2.24	1.12	2.42	1.15	1.89	0.92
CONTROL	before induction	30.46	8.47	21.2	10.96	1.17	0.52	1.23	0.52	1.72	0.98	2.46	1.15
	after induction	25.18	11.06	12.88	4.97	1.23	0.63	1.15	0.54	1.48	0.89	3.4	1.2

In the control condition, participants reported less positive affect after the interaction with the researcher ( $M = 25.18, SD = 11.06$ ) compared to before ( $M = 30.46, SD = 8.47; t(64) = 4.11, p < 0.001$ ), but they also experienced a decrease in negative affect after the induction ( $M = 12.88, SD = 4.97$ ) compared to before ( $M = 21.2, SD = 10.96; t(64) = 6.44, p < 0.001$ ; see Table 2.3).

More specifically, we found an decrease in distress after ( $M = 1.48, SD = 0.89$ ) the induction task compared to before ( $M = 1.72, SD = 0.98; t(64) = 2.34, p = 0.0225$ ), as well as a increased pride after ( $M = 3.4, SD = 1.2$ ) the induction compared to before ( $M = 2.46, SD = 1.15; t(64) = -6.78, p < 0.001$ ; see Table 2.3).

When comparing the affect data collected after induction between the control and the guilt conditions, participants reported higher positive affect ( $M = 4.82, SE = 1.65; t(65) = 2.93, p = 0.004$ ) and pride ( $M = 1.26, SE = 0.20; t(65) = 6.28, p < 0.001$ ) in the control condition. Moreover, they reported lower negative affect ( $M = -5.72, SE = 1.33; t(65) = -4.30, p < 0.001$ ), guilt ( $M = -1.42, SE = 0.174; t(65) = -8.34, p < 0.001$ ), distress ( $M = -0.52, SE = 0.19; t(65) = -2.70, p < 0.001$ ), shame ( $M = -0.37, SE = 0.18; t(65) = -2.17, p < 0.001$ ), and nervousness ( $M = -0.63, SE = 0.17; t(65) = -3.62, p < 0.001$ ) in the control condition compared to the guilt condition [*means and SE presented characterise the difference between the values in the control and the values in the guilt conditions*]. These results confirmed the effectiveness of the guilt induction method used: participants exposed to the guilt induction task reported higher levels of guilt and associated negative affect compared to those that were not.

### Principal Components Analysis

The Kaiser-Meyer-Olkin test (sampling adequacy: 0.507 in the control condition and 0.554 in the guilt condition) and the Bartlett's test of sphericity ( $\chi^2 = 41,361.120, p < 0.001$  in

the control condition and  $\chi^2 = 385,152.124$ ,  $p < 0.001$ ) confirmed suitability of the datasets to PCA. We reduced the data using a Varimax rotation, allocating AUs to components based on the highest factor loadings, whilst ignoring weak factor scores ( $<0.30$ ; Table 2.4).

**Table 2.4. Factor loadings of the Principal Component Analysis (PCA) – Guilt condition**

<i>Guilt condition</i>					
<b>All participants</b>					
AUs loaded	Rotated Component Matrix				
	Component				
	1	2	3	4	5
AU1	<b>0.951</b>	0.059	0.040	0.132	-0.006
AU2	<b>0.945</b>	0.062	0.054	0.078	-0.045
AU5	<b>0.379</b>	0.034	-0.065	-0.354	0.218
AU17	0.040	<b>0.762</b>	-0.038	0.087	-0.052
AU24	0.028	<b>0.759</b>	0.128	-0.063	0.006
AU18	0.046	<b>0.612</b>	-0.103	0.129	0.064
AU12	-0.007	-0.125	<b>0.738</b>	0.001	-0.033
AU7	0.044	0.073	<b>0.688</b>	-0.051	-0.133
AU14	-0.053	0.453	<b>0.477</b>	-0.069	-0.037
AU64	-0.125	0.035	0.141	<b>0.664</b>	-0.064
AU54	0.055	0.048	-0.056	<b>0.618</b>	-0.009
AU4	0.184	0.156	-0.119	<b>0.508</b>	0.052
FaceTouch	0.238	-0.188	-0.024	<b>0.312</b>	0.016
Guilt2	0.043	-0.024	-0.091	-0.049	<b>0.769</b>
NeckTouch	-0.033	0.026	-0.067	-0.041	<b>0.576</b>
AU10_11	0.073	-0.015	0.509	0.165	<b>0.511</b>

*Extraction Method: Principal Component Analysis; Rotation Method: Varimax with Kaiser Normalization. Numbers in bold indicate the higher loading for each AU and the contribution to a specific factor*

Examining the AU/ADs produced during the guilt induction task (guilt condition), we found that the PCA produced a 5-factor solution (see Table 2.4 for details). Felt guilt clustered with AU10 (Upper Lip Raiser) and Neck Touch in the fifth factor, which explained 7.51% of the variance. We found no significant relationship between self-reported shame or pride and AU10 (shame:  $\beta = 0.00045$ ; SE = 0.00026;  $p = 0.081$ ; pride:  $\beta = -0.00011$ ; SE = 0.00022;  $p = 0.63$ ) or Neck Touch (shame:  $\beta = 0.0016$ ; SE = 0.0014;  $p = 0.26$ ; pride:  $\beta = -0.0015$ ; SE = 0.0012;  $p = 0.21$ ).

Examining the AU/ADs produced during the pride induction task (control condition), the PCA also produced a 5-factor solution (see Table 2.5 for details); here, felt guilt clustered with AU62 (Look Right) in the fourth factor, explaining 9.84% of the variance. When the level of felt guilt is low, the reported emotion clustered with movements that were not significantly produced when felt guilt was high (guilt condition).

**Table 2.5. Factor loadings of the Principal Component Analysis (PCA) – Control condition**

<i>Control condition</i>					
<b>All participants</b>					
AUs loaded	Rotated Component Matrix				
	Component				
	1	2	3	4	5
AU1	<b>0.989</b>	0.004	0.030	-0.037	0.004
AU2	<b>0.987</b>	0.016	0.032	-0.035	0.023
AU57	-0.011	<b>0.764</b>	0.013	-0.112	-0.083
AU64	-0.061	<b>-0.680</b>	0.021	-0.153	0.087
AU51	-0.086	<b>0.474</b>	-0.072	-0.310	0.409
AU12	-0.049	-0.019	<b>0.796</b>	0.091	-0.047
AU14	0.100	0.008	<b>0.739</b>	-0.136	0.157
AU62	0.068	-0.043	0.024	<b>0.775</b>	0.020
Guilt2	-0.090	-0.016	-0.027	<b>0.530</b>	-0.028
AU61	0.059	-0.319	0.357	<b>-0.435</b>	-0.203
AU54	0.050	-0.166	0.100	0.078	<b>0.887</b>

*Extraction Method: Principal Component Analysis; Rotation Method: Varimax with Kaiser Normalization.*

*Numbers in bold indicate the higher loading for each AU and the contribution to a specific factor*

We found that participants produced similar rates of AU10 in both the guilt ( $M = 2.32$ ;  $SD = 2.19$ ) and control ( $M = 1.92$ ;  $SD = 3.06$ ) conditions ( $t(126) = -0.858$ ;  $p = 0.393$ ). They also produced similar rates of AU62 in the guilt ( $M = 6.36$ ;  $SD = 3.78$ ) and control ( $M = 5.68$ ;  $SD = 5.76$ ) conditions ( $t(108.7) = -0.793$ ;  $p = 0.430$ ). However, participants produced higher rates of neck touch in the guilt ( $M = 4.17$ ;  $SD = 22.45$ ) condition than in the control conditions ( $M = 0$ ;  $SD = 0$ ;  $U = 2,336$ ,  $p = 0.015$ ).





**Figure 2.2. Images of guilty expressions taken from videos.**

a) An expression typical of guilt production, AUs 1+2+5+10+12+25+26+**Neck Touch**; b) an expression typical of guilt perception, AUs 1+4+10+24+**Face Touch**). The individuals have given written informed consent (as outlined in PLOS consent form) to publish these case details.

### Population differences

To examine population differences we replicated the above analysis on the two populations, in each condition, using the same AUs identified in the general factor analysis. In the guilt condition, we found that felt guilt clustered with Neck Touch and AU5 (Upper Lid Raiser) in Population 1 (explaining 8.69% of variance) (predominantly of European origin), and with Neck Touch + AU 10 (Upper Lip Raiser) in Population 2 (explaining 8.34% of variance) (predominantly of East-Asian origin).

### Discussion

This first study aimed at eliciting an emotional response associated with the experience of guilt. We identified a pattern of non-verbal movements produced more when experiencing guilt: people raised their upper lip (AU10 Upper Lip Raiser) and touched their neck (Neck Touching). Those specific behaviours were not associated with self-reported feelings of shame or pride and were only associated with a higher level of felt guilt. Previous research that identified behavioural displays associated with embarrassment (gaze down, controlled smiles, gaze shifts, and face touches; Keltner, 1995), shame (head and gaze down; Izard, 1977; Keltner, 1995; Lewis et al., 1992) and pride (expanded posture, head tilted back, low

intensity non-Duchenne smile; Tracy & Robins, 2008) did not report AU10 and Neck Touch as part of those displays. Moreover, Neck Touch was produced significantly more in association with the experience of guilt and was not produced in the control condition. Therefore, Neck Touch seems the most specific to the experience of self-reported guilt that the participants reported in our study. We found little difference between the populations studied here, but participants from Population 1 (European origins) displayed neck touching with eye-widening (AU5 Upper Lid Raise), instead of upper lip movement. Little difference between the two populations is not enough to claim that those patterns are universal, but it seems to indicate that the behavioural expressions found here associated with guilt are present in more than one culture. This is the first study to identify a potential pattern of movements associated with felt guilt.

To determine whether these movements were also identified as guilt by observers, we conducted a second study. We asked new participants to examine the videos of participants from this first study during the induction of guilt (Figure 2.1:4) and rate them for emotion. We also asked these new participants to identify specific times when they thought they could see these specific emotions. This study aimed at identifying which specific facial movements were most closely associated with the perception of guilt.

### 2.3. Study 2 – Judgement of Guilt

#### Methods

##### Participants

One hundred and fourteen participants (hereafter referred to as “judges”) were recruited for this second study (82 females;  $M_{age} = 29.96$ ,  $SD = 11.48$ ; see Table 2.5). As in Study 1, judges were recruited based on an opportunistic sampling method and were all UK residents at the time of the experiment (but included both UK and non-UK nationals). All of them received either course credit (if a student) or £5 for their time. The whole experiment lasted 35 minutes on average. Our sample consisted of three distinct populations. Population 1 consisted of predominantly European University students and staff ( $N = 68$ ); Population 2 mainly consisted of overseas students, of predominantly Asian origin ( $N = 26$ ); and Population 3 consisted of overseas students of predominantly African/Caribbean origin ( $N = 20$ ); for a breakdown, see Table 2.5). Participants with European nationalities had stayed in the UK (mean = 25.60 years,  $SD = 19.87$ ) for significantly longer than participants with East-Asian nationalities (mean = 4.41 years,  $SD = 7.23$ ), and participants with African/Caribbean nationalities (mean = 1.38 years,  $SD = 3.53$ ;  $F(2, 67) = 29.8$ ,  $p < 0.001$ ). Most East Asian participants had recently arrived in the UK to study at the University (see Table 2.6 for details). The project has been reviewed and approved by the Science Faculty Ethics Committee (SFEC) from the University of Portsmouth (Appendix 3). Each participant signed an informed consent form granting authorisation for the use of the data for research purposes.

Table 2.6. Participants demographic information (Study2)

Nationality	Count	Average time spent in UK
<b>European</b>	<b>68</b>	<b>25.02</b>
American	2	1.36
American and Dominican	1	0.75
British	46	33.63
Canadian	1	0.08
Dominican & Spanish	1	0.07
Dutch	2	0.08
English & French	1	26.00
French	1	70.58
German	1	0.08
Italian	2	1.92
Polish	3	8.69
Romanian	2	2.00
Spanish	3	0.27
UK and Estonian	1	19.00
Ukrainian	1	0.08
<b>Asian</b>	<b>26</b>	<b>2.51</b>
Bangladeshi	1	0.33
British	1	10.16
Canadian	1	0.00
Chinese	8	0.32
Chinese & BNO	1	8.00
Chinese Hong Kong	1	7.00
French	1	4.00
Hong Kong China	1	0.04
Hong Kong Citizen, British National (Overseas)	1	11.00
Indonesian	3	0.75
Japanese	2	7.75
Malaysian	2	0.08
Thai	2	0.09
Vietnamese	1	4.00
<b>African/Caribbean</b>	<b>20</b>	<b>1.44</b>
Bolivian & Canadian	1	0.08
Botswana	1	2.70
British & South African	1	16.58
Dominican	13	0.59
Dominican & Spanish	1	1.00
Ecuadorian	1	0.08
Tanzanian	1	0.58
Venezuelan, United States Citizen	1	0.08
<b>Grand Total</b>	<b>114</b>	<b>15.75</b>

*N.B. Time spent in the UK presented in years; participants self-identified as European” or*

*“Asian” and their nationalities are presented in this table.*

**General Procedure**

At the start, participants were given general instructions regarding the experiment and written consent was obtained. Judges were originally told that this study had a different aim - to assess their abilities to detect facial expressions of emotion. Following this, the rest of the instructions and tasks were displayed on a computer through Qualtrics Survey Software (Qualtrics, 2012). Firstly, judges were asked to complete a personality questionnaire, the Guilt and Shame Proneness Scale (GASP, Cohen et al., 2011; order of questions was randomised between subjects) to assess their propensity to feel guilty and test whether this influenced their perception of guilt on the face. Secondly, judges were asked to watch 20 consecutive videos of faces, and make a series of ratings regarding the emotional state of the stimulus individual (see below for more detail). Finally, the judges were informed about the true aim of the experiment and debriefed. The experiment was presented on desktop computers in one of the laboratories available at the University. Judges sat in front of the computer, the screen situated approximately 60cm away from their faces (face stimuli visual angle: 10° x 14°). The Qualtrics survey was presented in full-screen mode; videos were uploaded on Youtube, on a private account, and presented on Qualtrics as an embedded file. Judges had the opportunity to watch each video as many time as they wished to and they could view it full screen. They could not slow down the video or watch it frame by frame but could scroll through the video.

**Stimuli**

All experimental video stimuli were taken from Study 1 (participants experiencing the guilt induction task, in the guilt condition). Of the 64 participants allocated to the guilt condition in Study 1, we used 57 individuals (seven participants were omitted for spending 50% of the time or more out of sight). Control video stimuli were also taken from Study 1 (participants from the control condition). For this, 12 individuals were chosen randomly. We wanted to include non-guilt videos to test the ability of participants to identify guilt on a face.

However, due to time constraint, we did not want to present more than 20 videos to each participant. Moreover, financial restrictions allowed us to recruit only 114 participants. Taken both time constraint and financial limitation into account, we had to present fewer control videos than guilt videos to each participant. All 69 videos (57 guilt and 12 control) were clipped to 30-90 seconds and audio was removed. The stimuli were generally centred in the video but participants in Study 1 were free to move their head and body.

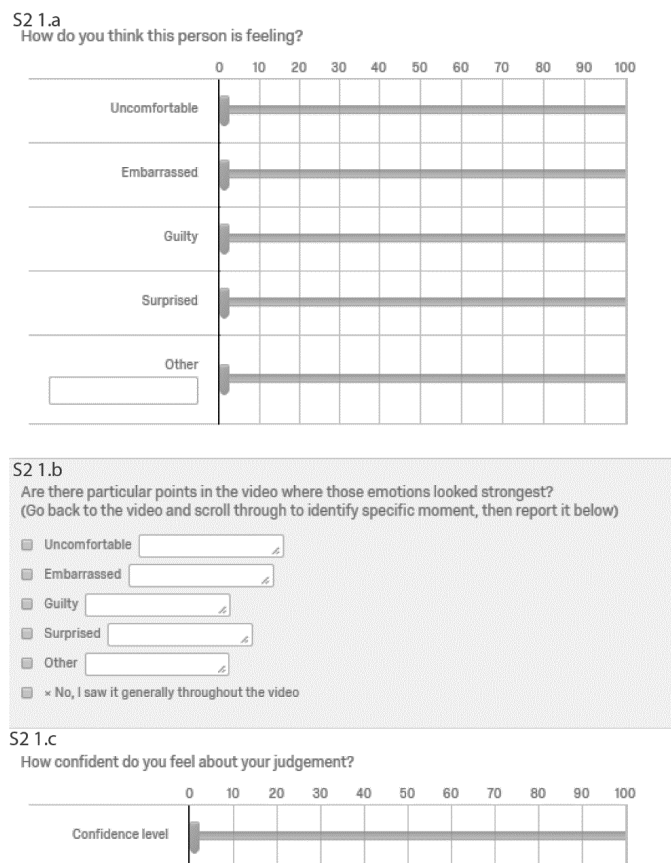
### **Guilt Judgements**

Each judge watched 20 videos in succession, 16 guilt videos and 4 control videos, out of the 69 videos selected for this study. Thus, each judge saw 16 guilt videos selected randomly out of the 57 videos available. Every judge saw a different series of guilt videos but at least 20 judges saw each video. The videos presented were randomised and counterbalanced for each participant. Before viewing the videos, judges were provided with the contextual information 'that the individual in the video had just been told they had wiped some important information from a USB flash drive'. The same contextual information was provided for all videos, guilt and control. While watching each video or right after viewing, judges were required to indicate how they thought the individual was feeling overall, using a sliding-scale (from 0-100%) for the five following emotional states: "uncomfortable", "embarrassed", "guilty", "surprised", and "other" (see Figure 2.3). Those judgements were collected for the entire video, as a measure of the different emotional states the individual in the video seemed to experience. The five emotional states were selected based on the results from Study 1, using the AUs produced during the experimental induction and identifying facial expressions associated with specific emotional states identified in previous research (AUs indicative of these 4 emotional states were identified on participants' faces; Ekman, 1971; Ekman & Friesen, 1969).

In addition to the sliding-scale rating, judges were encouraged to report any instances of emotion, i.e. any frame within the video where the emotion occurred (hereafter, a pinpoint), allowing for their judgements to be localised to an exact time point. They could have reported that overall the individual in the video experienced 20% of discomfort; this gave them the opportunity to indicate when exactly in the video was the individual experiencing discomfort. They were encouraged to report times when the indicated emotion was the most clearly expressed on the face (i.e., apexes of emotional expressions). To do so, participants could stop the videos whenever they wanted, watch the video multiple times, and even slow down the videos. Judges could not report a specific frame in the video due to the format of the stimuli, but they could report specific time (min:sec). Judges could make multiple pinpoints for multiple emotions, and multiple pinpoints per emotion. For example, they could report that in a video, the individual appears 50% guilty at 15 and 25 seconds in the video; or a judge could provide us with the information that an individual in a given video appears 50% embarrassed and 10% surprised at 35 seconds in the video, and 30% guilty at 40 seconds in the video.

When looking at the raw data, 623 instances of guilt were identified by all judges. This gives us a gross overview of the pinpoints reported as some of these instances might actually be the same pinpoint (or unique instance), as multiple judges might have reported the same specific time. Moreover, 1,077 instances of surprise were reported, as well as 825 instances of discomfort and 676 instances of embarrassment. Judges seemed able to conceptually differentiate between those four emotional states as very few overlaps were made between them (see the “*Descriptive analysis*” sub-section in Results for details on guilt pinpoints used in further analysis). We decided to look more in-depth at the facial movements identified during instances of surprise as Study 1’s design included a potentially surprising and unexpected element (i.e., being told they did something wrong unexpectedly). It would have been beyond the scope of this article to examine facial movements associated with

discomfort and embarrassment pinpoints. Moreover, judges seemed able to conceptually differentiate between those four emotional states as very few overlaps were made between them (see *Descriptive* in Results for details on guilt pinpoints).



**Figure 2.3. Judgement ratings.**

For each video, judges were asked to report **a.** how they thought the individual in the video was feeling using 5 sliding scales (from 0-100%); **b.** by indicating when they thoughts the emotion was most clearly expressed on the face; and finally **c.** they were asked to indicate how confident they were about their judgement.



**Compiling the dataset**

**Guilt.** Before analysis, the judgement data collected was combined with the FACS data produced in Study 1. The judges in this study reported 403 unique instances of guilt across the guilt videos and 67 unique instances across control videos, as identified by time-specific pinpoints on the video. We allowed for 0.5 seconds (or 12 video frames) of error around pinpoints, providing us with one second of video data per pinpoint in which judged guilt could have occurred. These pinpoints were synchronised with the FACS coding of the videos, to match judged guilt with any possible facial movements. We created these windows as the actual pinpoints reported by the judges were lacking precision; when synchronising the pinpoints with the FACS coding, we reported pinpoints in the middle of the second identified. For instance, a pinpoint identified by judges at 0:05 would be reported in the FACS coding at 0:05:500. The 1-second window allowed us to capture the movements they perceived as reporting guilt. Moreover, as genuine expressions have been shown to have onsets ranging from 0.50 to 0.70 sec (Cohn & Schmidt, 2003; K. L. Schmidt et al., 2009), we thought that creating a 1-second window around the identified pinpoints would allow us to capture the facial movements identified as conveying guilt by the judges. We were not interested in capturing the unfolding of the entire expressions associated with guilty feeling, from onset to offset; rather we wanted to explore facial movements people associate with guilt. Finally, any video data that occurred outside of these pinpoints (i.e. any part of the video that was not judged as guilty by any judge) was removed, providing us with a reduced dataset containing only judged guilt video frames. The creation of the pinpoints and removing all frames occurring outside the pinpoints resulted in 8,934 video frames of FACS data (present/absence of AU/ADs) from the guilt videos and 850 video frames from the control videos. This step was conducted to focus our data more so on facial movements the judges were considering as guilty, and to reduce noise in the dataset. All the selected frames were retained for further analysis.

As in Study 1, any AU/AD that was rarer than the average in this new dataset were excluded. We used a binomial exact test as criteria for exclusion - if any AU/AD were produced by fewer participants than the calculated criteria (here, the criteria calculated with the binomial test was a minimum of 35 participants in the guilt condition and 8 in the control condition), this AU/AD was not explored further in an attempt to maintain a robust dataset. We also ran a correlation between each AU and the judged guilt. This again reduced our dataset down to 14 AU/ADs in the guilt videos and 8 AU/ADs in the control videos, out of a possible 39 observed in our data (see Table 2.2 for details).

**Surprised.** The same procedure was used to select AU/ADs reliably associated with judged surprise. Judges identified 427 unique instances of surprise across guilt videos and 32 across control videos. We compiled those data with the FACS coding for each video; we then created the 1-sec window around the surprised pinpoints and removed any frame outside those windows. We selected AU/ADs using a binomial exact test (criteria: 33 participants in guilt videos and 8 in control videos) and running correlations between each AUs and judged surprise. We selected 10 AU/ADs across guilt videos and 8 AU/ADs across control videos.

### Statistical Analysis

**Guilt.** First, to examine the judges' ability to accurately perceive guilt on a face, we ran a Pearson's correlation between the reported felt-guilt of each participant and the averaged judged guilt per participant (i.e., video).

Second, to identify AU/ADs of interest, we ran two factor analyses (Principal Component Analysis with Varimax rotation, PCA), one for each type of videos (guilt or control), on the frequency data of the selected AU/ADs. PCAs were conducted on frame-by-frame datasets to produce factors consisting of temporally clustered AU/ADs as done previously (Stratou et

al., 2017). We added the judged guilt to both PCAs in order to identify movements that were produced more often when guilt was identified. PCA was conducted in SPSS version 24 (IBM, 2016).

Third, to understand how the movements were temporally produced, we looked at the production of AU/ADs over time. The pinpoints created one-second windows (from -0.5 second to 0.5 seconds, or -12 video frames to +12 video frames, around judged guilt), allowing us to quantify facial movement change over time. To do this, we conducted a curvilinear model analysis, to see which linear or curvilinear pattern any potential facial movements of guilt most closely followed. This curvilinear model analysis was conducted in SPSS version 24 (IBM, 2016).

Finally, as our sample was made up of distinct populations (Population 1, 2, and 3), two of which matched the populations of the participants in Study 1, we again used this as an opportunity to assess for population differences in the perception of guilt, in guilt videos. To do this we repeated the same steps as above: we divided our dataset between the guilt videos from those of predominantly European origin (Study 1: Population 1) and the guilt videos of those of predominantly Asian origin (Study 1: Population 2). We retained all populations of judges, but we investigated whether or not their reported cultural identity matched with that of the stimulus individual (congruent 1, or not 0), using a generalised linear mixed model approach (GLMM). Judged guilt (0-100%) of the stimulus individual was set as our response variable. As our dataset considered ratings of videos by multiple judges, we include the identity of the judge, and the stimulus individual, as random factors in the model to control for pseudo-replication (Hurlbert, 1984; Waller, Warmelink, Liebal, Micheletta, & Slocombe, 2013). We fitted the GLMM using the function `lmer` provided by the packages `lme4` and `lmerTest` for RStudio Version 0.99 for R 287 version 3.1.3 (Bates et al., 2014). We fitted the GLMM with a Gaussian error structure and estimated using REML.

**Surprise.** To identify AU/ADs of interest, we ran again two factor analyses (PCA with Varimax rotation), one for each type of video (guilt or control), on the frequency data of the selected AU/ADs. PCAs were conducted on frame-by-frame datasets and we added the judged surprise to both PCAs in order to identify movements that were produced frequently more often when surprise was identified. PCA was conducted in SPSS version 24 (IBM, 2016).

## Results

### Descriptive statistics

Overall, judges attributed a higher level of guilt to participants in the guilt videos ( $M = 35.65$ ,  $SD = 9.46$ ) than to participants in control videos ( $M = 19.16$ ,  $SD = 7.47$ ;  $t(19.246) = -6.61$ ,  $p < 0.001$ ,  $d = 1.49$ ). They also attributed higher level of surprise to the guilt videos ( $M = 42.46$ ,  $SD = 16.18$ ) than to the control videos ( $M = 21.48$ ,  $SD = 12.93$ ;  $t(2.051) = -4.88$ ,  $p < 0.001$ ,  $d = 1.20$ ).

The judges reported 403 instances of guilt across the guilt videos, with an average of seven pinpoints per video, and 36 instances across the control videos. In 40 of those instances (10% of the total amount of guilt pinpoints identified across all videos), guilt was associated with one other emotion (guilt was associated with embarrassment in 45% of these 58 occurrences, with discomfort for 47.5% and surprise for 7%; see S2 Table 2.7).

Judges made reliable ratings regarding the level of felt guilt: we found a positive correlation between the averaged judged guilt per individual and the individual self-reported guilt ( $r=0.465$ ,  $n=69$ ,  $p < 0.001$ ).

**Table 2.7. Descriptive overview of guilt pinpointed by judges**

	Guilt with another emotion	Guilt occurring with (details)			Guilt alone	<i>Grand Total</i>
		Discomfort	Embarrassment	Surprise		
Occurrences	40	18 (45%)	19 (47.5%)	3 (7%)	363	403
Percentages	10%				90%	

*N.B. percentages in bracket indicate the proportion out of 58 pinpointed when guilt was reported with another emotion*

### Principal Components Analysis

**Guilt.** Examining the AU/ADs, we found that in the guilt video, judged guilt clustered with AU4 (Brow Lowerer) and Face Touch, in a factor explaining 10.59% of the variance.

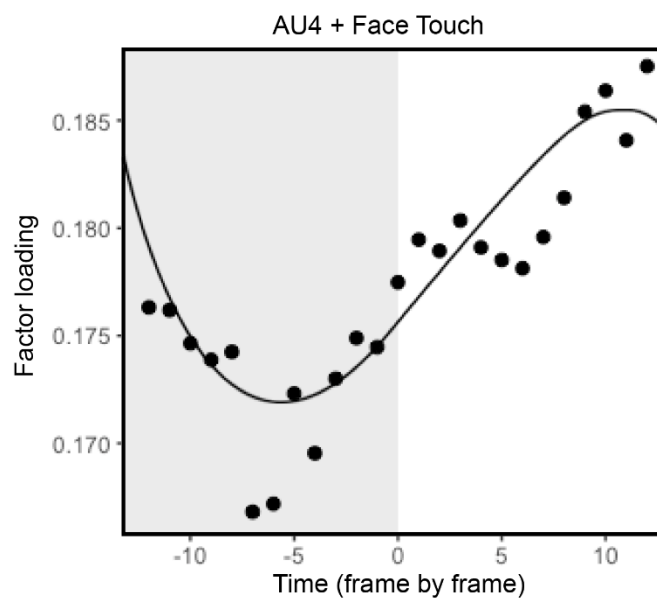
In the control videos, judged guilt clustered with AU37 (Lip Wipe) and AD59 (Up Down Head Shake), movements that were not associated with a higher level of judged guilt (guilt condition).

**Surprise.** In guilt videos, judged surprise clustered with AU10 (Upper Lip Puller), explaining 10.33% of the variance.

In control videos, judged surprise clustered with AU1 (Inner Brow Raise), AU2 (Outer Brow Raise), AU55 (Tilt Head Left) and AU61 (Look Left) in one factor explaining 29.39% of the variance; it also clustered with AU28 (Lip Suck) and AU57 (Head Forward) in a second factor explaining 15.59% of variance.

### Facial movements of guilt over time

The curvilinear analysis showed that when looking at the production of AU4 + Face Touch over a 1 second (25 frames) window around pinpoint of judged guilt, the data best fit to a cubic pattern ( $r^2 = 0.940$ ,  $F(3,21) = 110.611$ ,  $p < 0.001$ ). This potential facial movement of guilt appears to increase in intensity over time, and continue to increase after the signal is recognisable (see Figure 2.4).



**Figure 2.4. Representation of the temporal production of facial movements.**

The time course of the production of the PCA factor Factor 2C is presented: Factor 2C (AU4+Face touching) was increasingly produced from the pinpoint to apex 300ms later. Grey area represents pre-guilt period; white area represents post-guilt periods. Time is represented by the number of frame, 12 frames = 500 ms.

### Population differences

We conducted the following analysis on the two populations using the same 14 AUs as above (guilt videos). In those PCAs, judged guilt clustered with AU4 (Brow Lowerer) and Face Touch in both populations (European and Asian).

Congruence between the ethnicity of the judge and the individual in the video had no effect on judgement of either the European or Asian participants ( $\beta = 2.45$ ,  $SE = 4.53$ ,  $p = 0.590$  and  $\beta = -1.94$ ;  $SE = 5.23$ ;  $p = 0.711$ ).

## **Discussion**

This study aimed to identify which facial movements were perceived as guilt when guilt was induced in a laboratory experiment. We found that when judges gave a higher rating of guilt, people were frowning (AU4 Brow Lowerer) while touching their face (Face Touching). We used instances when judges reported seeing guilt to create 1s-window of interest and conduct our analysis only on those time windows of guilt. Doing this, we identified facial movements reliably associated with the perceived guilty expression. Judges reported other emotions at the same time as guilt in only 14% of the pinpoints of guilt identified. Moreover, pinpoints of guilt revealed specific facial movements that were not present either in control videos or in association with pinpoints of surprise. This made us fairly confident that the facial expressions identified were associated with the experience (perception) of guilt.

## 2.4. General Discussion

In two studies, we aimed to identify facial movements and behavioural displays associated with the experience of guilt in humans. In the first study, we examined the production of guilt using a novel induction technique. In the second study, we examined whether others perceived guilt from the face of those experiencing guilt. We used an extensive, bottom-up coding scheme to identify facial patterns associated with the experience (production and perception) of guilt as part of a dynamic sequence of behaviour.

We found a positive relationship between the level of self-reported guilt and the extent this individual was judged as feeling guilty by others. This supports the idea that guilt could have evolved as an observable phenomenon with a potential communicative social function. The patterns identified in this experiment showed slight variation between what people do when feeling guilty and what people see when identifying guilt. Our first study showed that guilt was associated with upper lip movement and neck touching (AU10 Upper Lip Raiser; Ekman & Friesen, 1978), with Neck Touch being the most specific non-verbal signal associated with guilt. Our second study showed that the identification of guilt in others was associated with frowning and face touching (AU4 Brow Lowerer; Ekman & Friesen, 1978). We found little population differences in the production of a guilt non-verbal signal. Participants from European populations displayed eye-widening (AU5 Upper Lid Raise) and neck touching, instead of upper lip movement (AU10) and neck touching. We did not find any difference in the non-verbal signal perceived as guilt, because all participants used the frowning (AU4) and face touching signal to detect guilty feelings.

Though we did not find a direct association between felt guilt and the perception of guilt, we found that both felt guilt and perceived guilt were associated with self-directed behaviour. Moreover, when looking at the non-verbal signals produced during the



experience of guilt, we found that guilty individuals produced more frown (AU4, Brow Lowerer; Ekman & Friesen, 1978) and face touching (i.e., facial signals used by observers) than individuals in the control condition. Thus, while there is no direct correspondence between what is produced when guilty, and what looks like guilt to others, there is a similarity. It is possible that the nature of the perception task, involving verbal conceptualisation of an emotional state, influenced observers; the non-verbal signals perceived here as indicating guilt might not be the signals one would spontaneously respond to in a spontaneous interaction. More recent conceptualisations of emotional experiences (Barrett, 2017a, 2017b; Fridlund, 1994, 2017; Scherer et al., 2013) argue for a less universal and omnipotent link between the experience of an emotion and behavioural outcomes. In an emotional context, multiple systems will be triggered (e.g., cognitive processes, physiological systems, motor expressions; Scherer et al., 2013), leading to multiple behavioural outcomes (e.g. non-verbal signals), one of which might be used by observer when responding to the situation (Scherer et al., 2013). As such, an individual feeling guilty might produce multiple non-verbal signals, one of which will be more strongly associated with the subjective, constructed, feeling of guilt; an observer might perceive those non-verbal signals and rely mainly on specific ones to interpret the emotional state of the guilty individual.

Using a bottom-up methodology allowed us not only to approach our question without any a priori assumptions regarding the results, but it also increased the likelihood that the movements identified in our studies are associated with the experience of guilt and no other secondary moral emotion. Indeed, the “guilt” pinpoints identified by the judges (Study 2) were mainly instances of identification of guilt alone, with only 14% of the total number of guilt pinpoints associated with more than one emotion (see Table 2). This allowed us to focus our analysis on facial movements associated with the experience of guilt only. Moreover, even though guilt is often mistaken for embarrassment or shame, the embarrassed display

has been characterised by the joint production of gaze down, controlled smiles, head turns, gaze shifts, face touches (Keltner, 1995), and the occasional blushing (de Jong & Dijk, 2013); and the typical face of shame was described with head and gaze movements down (Izard, 1977; Keltner, 1995; Lewis et al., 1992). None of the movements we found in relation to the expressions of guilt were found associated with those other negative self-conscious emotions. During the AU selection process, most facial movements associated with either embarrassment or shame were discarded from further analysis, with the only exception of face touching. Face touch can emphasise embarrassment displays, but it is not necessary for the identification of embarrassment (Keltner, 1995). A previous study suggested a link between blushing and admission of guilt (De Jong, Peters, & De Cremer, 2003); combining FACS analysis with thermal imaging techniques might have revealed changes in facial temperature in guilty participants, which could be unconsciously used by observers in their judgments. Moreover, when looking at the facial signals associated with the pinpoints of judged surprise (judges reported the highest number of instances for this emotion), we did not find any of the non-verbal signals of guilt. It would thus appear that the pattern of facial movements identified in our studies is specific to the emotion of self-reported guilt experienced by our participants.

Both the production and perception of guilt was associated with self-directed behaviour (i.e., scratching, neck or face touching), which are often classified as displacement behaviours, and are defined as a group of behaviours that appear irrelevant to the situation in which they are displayed, but can gain communicative value over time (Troisi, 2002). The production of such behaviours has been shown to increase in stressful, negative, situations (Mohiyeddini, Bauer, & Semple, 2013; Mohiyeddini & Semple, 2013). Self-directed behaviours may be used when individuals try to distance and protect themselves from an unpleasant situation, acting as a short-term diversion of attention, which could, in turn, reduce the negative feeling associated to the situation at hand (Chance, 1962; Mohiyeddini

et al., 2013; Sgoifo et al., 2003). Self-directed behaviour could thus help regulate the level of stress associated with emotionally challenging situations (Mohiyeddini & Semple, 2013), such as the guilt induction experienced by our participants in Study 1. Indeed, some studies have shown that self-directed behaviours are common in situations such as embarrassment (Keltner, 1995), discomfort (Ekman & Friesen, 1969), and anxiety and guilt (Ekman & Friesen, 1972), which focussed on hand movements and found a correlation between the production of self-adaptor (i.e., scratching) and anxiety and guilt feelings. In our study, we found that the experience of guilt was associated with self-directed behaviours (neck and face touching), which appears to be in line with previous research. However, the production of self-directed behaviours could be due to the experimental design: participants were seated at a table, in front of a computer. Such setup could have elicited those movements, even though participants in the control condition, also seated at a computer, did not display as many self-directed behaviours.

The bottom-up methodology also diverges from previous research examining the facial display of guilt, which is why we may have found a more concrete candidate for the display of guilt. One notable previous study used a literature-based conceptualisation of the experience of guilt to present three candidates' displays to their participants (Keltner & Buswell, 1996). In that study, using a top-down approach, the participants were presented with displays selected based on previous literature, which associated the experience of guilt with the experience of self-contempt, sympathy, and pain. The authors tested whether their conceptualisation of guilt accurately described a facial display associated with the experience of the emotion. The results were not conclusive as the candidates' displays were more often associated with emotions other than guilt (Keltner & Buswell, 1996). A more recent study associated the experience of guilty feeling with increased skin conductance and gaze avoidance (Yu, Duan, & Zhou, 2017). We did not find gaze avoidance (i.e. actively avoiding to look in another person's direction) to be part of the non-verbal signal of guilt, even though

participants in the guilt condition looked down more than participants in the control condition. Yet, this could be due to our experimental design: participants in the guilt condition might have been looking down at the laptop more than people in the control condition. So, it is unclear in our design whether guilty participants avoided eye-contact or focused on an object associated to their wrongdoing (the laptop could be incriminated for the deletion of data on the USB stick, removing the fault from them).

Guilt can be induced by various experiences and situations (i.e., interpersonal transgression, trust violation, private/public dimension); can be experienced at various degrees (i.e., trace to extreme); and can lead to multiple outcomes (i.e., make amends, get angry; Baumeister et al., 1994; Nelissen & Zeelenberg, 2009). Those different experiences of guilt might vary in their behavioural expressions, occasionally presented as movements associated with pain or self-contempt (Keltner & Buswell, 1996) or gaze avoidance (Yu et al., 2017), but with specific core movements different from the previous conceptualisation. Presenting dynamic stimuli and allowing for genuine production and identification of guilty expressions may have allowed us to identify a non-verbal signal commonly seen as guilt in humans, but it is also possible that our results are likewise limited to a specific context.

As our sample was composed of individuals from different populations, we conducted some cross-population analysis. We found that both populations produced very similar patterns of facial movements when experiencing guilt; however, participants from the European population group produced self-directed behaviour (neck touching) with eye-widening movement, instead of lip movement. The production of AU5 by European participants could be a population-specific accent (Chen & Jack, 2017; Jack et al., 2016) in the expression of guilt, but the results of these studies are only suggestive of population differences and are insufficient to claim that participants from different cultural backgrounds present variation in their display of guilt. To make such a claim, more in-depth cultural studies

are required, sampling participants from different countries, conducting the experiment in those countries and in the local language.

Here, we looked at the spontaneous production of genuine non-verbal expressions of emotion. Judges had to rely on genuine, dynamically presented facial expressions to recognise and rate emotions. Our experiments support a drive towards a new scientific culture, studying facial expressions using novel approaches removed from the dichotomous debate about nature vs nurture (Crivelli et al., 2016; Jack et al., 2016). Taken together, our two studies provide information regarding the production and the perception of guilt in each video. For each participant, we had a self-report of guilt, the facial movements produced at the time of the induction, and various points in time when judges identified guilt on participants' faces. We have identified a non-verbal signal of guilt that is recognised by receivers, suggesting an adaptive significance to feeling guilty and communicating this guilt to others. Such non-verbal signal of guilt could have evolved in humans as a strategy to communicate intentions (Fernandez-Dols, 2017), or to repair relationships, and may ultimately function to maintain strong cohesive social groups. Previous research extensively looked at the behavioural consequences of guilty feelings: it can promote directed action towards those who have been wronged (Cryder et al., 2012); it can reduce prejudice behaviours (Amodio et al., 2007); and can increase generosity (Ketelaar & Au, 2003). We focussed on the first reactions people have when realising they did something wrong and the guilty feelings emerge and were able to identify reliable candidates characterising the experience of self-reported guilt. Building on this, we conducted a study to investigate guilty people's propensity to repair the relationship, as well as the impact of a facial expression on the person wronged, i.e. the victim's reaction (Julle-Danière, Whitehouse, Vrij, Gustafsson, & Waller, in prep).

It is, however, important to remain cautious in the interpretation of our data, as our study also present limitations. Our results are valid in the given context of this experiment and for this specific induction of guilt. Different experiences of guilt will probably vary in their behavioural expressions, presenting maybe a core pattern of specific movements, but differing between situations. We also need to consider the fact that providing contextual information might have influenced the judges in their decisions. Additionally, the analyses conducted here might have biased the results; indeed, if the use of PCAs allowed us to conduct exploratory analyses, other more advances methods might have produced unskewed results (e.g., bootstrapping analysis; Snijders & Borgatti, 1999). Amother aspects of the procedure used here that could have biased the results is the disproportionate number of guilt and control videos viewed by judges (16 guilt; 4 control). This unequal split might have lead to a biased hit rate if the judges' responses were biased towards the "guilt" and the "none of the above" categories Moreover, our judgement study also presents some linguistic limitations. Even if there are differences in the appraisal and behavioural outcomes between shame and guilt, it has been previously shown that English speakers use "guilt" and "shame" interchangeably (Fessler, 2004). To assess the impact of context and assess the linguistic barrier, we conducted a follow-up study comparing the judgements made with and without contextual information provided (Julle-Danière, Whitehouse, Harris, et al., in prep; Chapter 3). By removing the contextual information, we hoped to gauge how the expression of guilt is then perceived when no verbal/written content needs to be understood first.

Additionally, to compare various judgement methodologies (emotion words vs action tendencies vs dimensions), we conducted another follow-up study (Chapter 4) to help us have a better understanding of how people conceptualise the facial expression produced when experiencing guilt, using different types of words and classification methodologies (forced-choice vs free labelling vs dimensions). In this way, we hoped to introduce more variability in the emotional judgements, looking at patterns of mislabelling of guilty displays. We showed that

people do not reliably identify a facial expression associated with guilt when using explicit emotion or action tendency labels, but do distinguish the expression of guilt from other emotional displays when using dimensional ratings (aroused-asleep; pleasant-unpleasant). Those two studies taken together (Chapters 3 & 4) emphasise the importance of context in understanding emotional displays associated with secondary emotions, as well as starting to shed light on the cognitive processes involved in interpreting them.

Furthermore, both our participant and judge samples were female-dominant (over 70% of people were females in both studies), which could have skewed our results if guilt is perceived differently in men and women (Clark, 1990). If guilt is categorised as embarrassment and shame as a display signalling lower status and submissiveness (Clark, 1990), it could be more easily perceived in women, as stated by previous research (Castelfranchi & Poggi, 1990; Keltner, 1995). Our samples did not allow us to compare between genders but it is something to keep in mind. More balanced samples could help solve this issue. Finally, we used participants with a European or Asian heritage, but with different nationalities, and potentially slightly different cultures and social norms. Participants could thus be identifying slightly different expressions of guilt. However, using broad cultural groups allowed us to identify a facial signal commonly seen as guilt in humans and is groundbreaking in the study of the expression of guilt. Acculturation questionnaires should be used in future studies to measure the cultural distinction between participants from different ethnical background.

Those are the first studies to look at the genuine expression of guilt and the perception of secondary emotion using dynamic stimuli. There were exploratory studies, using simple analysis and focussing on the behavioural signals associated with a guilt-inducing situation. We have however collected more extensive data; now that we identified a non-verbal signal associated with the experience of guilt, more in-depth analysis (such as a lens modelling, Scherer et al., 2013) would be an interesting step to further break down the mechanisms associated with guilt.

Future research would be needed to understand how temporal variations in expressivity affect the salience of the overall expression or the perception of the information. In our study, we tried to keep the rating of videos as simple as possible, leading us to suboptimal judgments. It would be interesting to conduct such experiments using coding software (such as BORIS, Friard & Gamba, 2016; or Interact, Mangold, 1998) to collect more detailed coding (e.g., temporally precise pinpoints, frame-by-frame coding). Future research should also explore how guilt expression may differ between cultures and situations and how other modalities such as body posture could impact the perception of such emotion (Martinez, Falvello, Aviezer, & Todorov, 2016). Finally, investigating the adaptive value of these expressions by considering how exactly they provide a benefit to the actors and receivers can shed new light about the evolutionary processes shaping them. Together, our results suggest that guilt is expressed on the face and communicates the experience of guilt to others through a non-verbal signal.



### 3. *Guilt outside of context*<sup>4</sup>

## The impact of contextual information on the perception of a facial signal of guilt

### Abstract

Contextual information appears to be a key element in accurately understanding facial expressions. Guilt, a complex emotion with a potentially important social function of stimulating cooperative behaviours towards and from others, has been shown to be associated with facial signals associated with self-reported feelings as well as perceived feelings in others. Here, we investigated the relative importance of context in interpreting facial signals associated with secondary emotion, but also the universal aspect of perceived guilty signals. We conducted a replication study, presenting videos of guilty individuals with and without contextual information to a sample of US-based students. We replicated all our previous findings and we also showed that facial expression of guilt is better interpreted when the situational context is provided alongside the facial expression. It supports the trend for ecologically valid judgment research as a facial display is rarely occurring without context in real life.

**Keywords:** Facial expression, guilt, emotion, FACS, culture, self-directed behaviour

---

<sup>4</sup>Study conducted in collaboration with Christine Harris and Mingi Chung, Psychology Department, University of California, San Diego (USA).

### 3.1. Background

Humans are highly cooperative with both relatives and strangers (Tomasello, 2008), and the need for cooperation may have provided a powerful selection pressure behind many of the behaviours which we consider uniquely human. Guilt is an emotional and cognitive experience arising when someone feels that they did something wrong. It could have evolved due to its potential adaptive function, within social interaction, of prompting pro-social interactions towards and from others. It is classified as a moral, self-conscious emotion, along with pride, shame and embarrassment (Haidt, 2003), and is one of the most social, other-oriented emotions that people experience throughout life (J. Carroll, 1985; Cryder et al., 2012). Despite a growing literature on the social consequences of feeling guilty (Cryder et al., 2012; De Hooge et al., 2007; Ketelaar & Au, 2003), little is known about the behavioural mechanisms underpinning the social functions of guilt, specifically whether guilt is associated with a specific facial expression or nonverbal signal. Our previous study examined what people saw when identifying guilt on the face of another (Julle-Danière et al., under review; see Chapter 2). We induced guilt experimentally, eliciting patterns of movement that were associated with both the participants' subjective feelings of guilt and judges' impressions of their guilt (Julle-Danière et al., under review; see Chapter 2).

Previous research investigated the impact of context on judgement, by providing written contextual information (Fernandez-Dols, Sierra, & Ruiz-Belda, 1993); by showing a face surrounded by other faces, often presenting either congruent or incongruent expressions (Hess et al., 2016); or by presenting an expressive face on congruent or incongruent bodily postures (Aviezer et al., 2008). Those set of studies showed that in the presence of incongruent contextual information, the face was not as clearly associated with a specific emotion as it was when provided with congruent information or even in absence of context. These studies were conducted using sets of prototypical facial expressions of primary

emotions, which have previously been extensively studied with and without context. Numerous papers presented discrepancies in interpretations of those expressions (Crivelli et al., 2016, 2017; Jack et al., 2016), questioning the universality of the message conveyed. Less research has been conducted on facial expressions of secondary emotions (Keltner & Buswell, 1996), more idiosyncratic and context-dependent, and to our knowledge, none looked at the impact of context on their interpretation.

### **Present Investigation**

Here, we investigated whether judges living in the USA not only identified the same facial expressions as associated with the experience of guilt as UK judges (Julle-Danière et al., under review; see Chapter 2), but also whether they could reliably identify expressions associated with guilt in absence of context. I used the same methodology as Chapter 2 – Study 2, using the videos collected in Chapter 2 – Study 1. Only the perception study is described below.

## **3.2. Perception of guilt by a US-sample**

### **Participants**

One hundred and fifty-six participants (hereafter referred to as “judges”) were recruited for this replication study (134 female; mean age = 21.06, SD = 2.71). Participants were recruited online through the University of California, San Diego participant pool, as part of a larger project conducted by Christine Harris and Mingi Chung. All participants were US residents at the time of the experiment (including both US and non-US nationals). Participants received course credit for their time. Our sample consisted of two distinct populations. Population 1 consisted of predominantly American University students (N = 76, 72 female) and Population 2 consisted of overseas students, of predominantly Asian origin (N = 80, 62 female); for a breakdown, see Table

3.1. The project has been reviewed and approved by the Science Faculty Ethics Committee (SFEC) from the University of Portsmouth (Appendix 3).

**Table 3.1. Participants demographic information.**

<b>Nationality</b>	<b>Count</b>	<b>Average time spent in USA</b>
<b>Asian</b>	<b>80</b>	<b>10.60</b>
American	29	19.04
Canadian	3	8.67
Chinese	27	4.42
Hong Kong, British National (Overseas)	2	5.00
Indian	1	2.00
Indonesian	2	2.75
Irish	1	0.33
Korean	6	8.36
Macau	1	3.00
Malaysian	2	2.50
Taiwanese	1	5.00
Vietnamese	2	17.00
American/ Japanese	2	7.09
American/ Chinese	1	21.33
<b>American</b>	<b>76</b>	<b>19.92</b>
American	60	21.49
American/ Russian	2	16.75
Chaldean	1	23.00
European	1	21.00
German	1	1.83
Iranian	3	11.33
Jordanianian	1	18.00
Persian	2	6.00
American/ Israeli	2	15.38
American/ Bulgarian	1	13.33
Armenian/ Iranian	1	17.92
American/ Turkey	1	19.53
<i>Grand Total</i>	156	15.14

*N.B. Time spent in the USA presented in years*

## Procedure

The study was presented as “Can you see how I feel?” and judges were told that the purpose of the study was to look at their ability to detect facial expressions of emotions. Written consent was obtained for all judges using an online form and they received general instructions regarding the experiment, with specific instructions given on the computer screen in due time for each successive task. The experiment was presented on a computer using the Qualtrics survey software (Qualtrics, 2012) and was composed of two successive tasks. Judges were first asked to complete a personality questionnaire, the Guilt and Shame Proneness Scale (GASP; Cohen et al., 2011); questions were randomised automatically between subjects. Judges were then asked to watch 20 consecutive videos (see below for details). At the end of the study, all judges were fully debriefed and informed of the real goal of this study: investigating the facial expression of guilt.

## Stimuli

The videos used in this study are the same videos as the videos used in Julle-Danière et al’s Study 2 (under review; see Chapter 2 for details). Clips ranged from 30 seconds to 90 seconds. Due to time constraint, we did not want to present more than 20 videos to each participant: 16 guilt videos and 4 control videos. Control videos presented genuine, spontaneous facial expressions differing from the displays of guilt and were used to test the ability of participants to identify guilt on a face. Videos were randomised between judges ensuring a minimum of 20 judges watching each video. Half the sample (N = 78; 40 from Population 1) saw the videos with the following context provided for all the videos: “This person has just been told that they wiped important data from a USB stick”; the other half (N = 78; 36 from Population 1) of the sample viewed the videos without any contextual information. Judges were asked to provide a judgement regarding the emotional state.

## **Guilt Judgements**

For each of the 20 viewed videos, judges were asked to indicate how they thought the person in the video was feeling using a sliding-scale (from 0 to 100%) for the following five emotional states related to guilt: “uncomfortable”, “embarrassed”, “guilty”, “surprised”, and “other”. Judges also had the option for each emotional state to pinpoint the specific time in the videos when they thought the emotion was the strongest. They could give one or multiple time points per video when a given emotion was the strongest (see Chapter 2 for detailed procedure).

## **Compiling the dataset**

As significant differences were found between guilty and control videos in Chapter 2, we removed the control videos from the analysis. All following analysis was conducted on the 57 guilty videos only: of the 64 participants allocated to the guilt condition in Chapter 2 - Study 1, seven participants were omitted for spending 50% of the time or more out of sight. Each judge saw 16 guilt videos selected randomly out of the 57 videos available. Every judge saw a different series of guilt videos but at least 20 judges saw each video.

Before analysis, the judgement data collected (with or without contextual information) was combined with the FACS data produced in Study 1 (Julle-Danière et al., under review; see Chapter 2). When presented with a context, judges reported 203 instances of guilt across the guilt videos, as identified by time-specific pinpoints on the video. In the absence of contextual information, judges reported only 161 instances of guilt. We allowed for 0.5 seconds (or 12 video frames) of error around pinpoints, providing us with one second of video data per pinpoint in which judged guilt could have occurred. These pinpoints were synchronized with the FACS coding of the videos, to match judged guilt with any possible facial movements. Finally, any video data that occurred outside of these pinpoints (i.e. any part of the video that was not judged as guilty by any judge) was removed, providing us with

a reduced dataset containing only judged guilt video frames. This left us with 12782 video frames of FACS data (present/absence of AU/ADs) across conditions. This step was conducted to focus our data more so on facial movements the judges considered as guilt, and to reduce noise in the dataset.

Any AU/AD that was rarer than the average in these new datasets were excluded. We used binomial exact tests as criteria for exclusion - if any AU/AD were produced by fewer participants than the calculated criteria (here, the criteria given by the binomial test was a minimum of 33 participants in condition 1 and 30 participants in condition 2), this AU/AD was not explored further in attempt to maintain a robust dataset. This reduced our dataset down 13 AU/ADs across conditions, out of a possible 39 observed in our data (see Table 2.2 for details).

## **Statistical Analysis**

### **General analysis**

First, to examine the judges' ability to accurately perceive guilt on a face, we ran a Pearson's correlation between the reported felt-guilt of each participant and the averaged judged guilt per participant (i.e., video). We ran one correlation per condition (with or without context).

Second, to identify AU/ADs of interest, we ran a factor analysis (Principal Component Analysis with Varimax rotation, PCA) on the frequency data of the 13 AU/ADs. PCA was conducted on frame-by-frame data to produce factors consisting of temporally clustered AU/ADs. PCA was conducted on frame-by-frame datasets to produce factors consisting of temporally clustered AU/ADs as done previously (Stratou et al., 2017). We added the judged

guilt to the PCA in order to identify movements that were produced more often when guilt was identified. PCA was conducted in SPSS version 24 (IBM, 2016).

Finally, we ran a generalised linear mixed model approach (GLMM) to assess the impact of the condition and cultural congruency on the perception of guilt. Judged guilt (0-100%) of the stimulus individual was set as our response variable, and the condition (with or without context), the cultural congruency (congruent 1, or not 0), and the frequencies of AUs clustering with judged guilt identified in the PCA were set as our predictor variable. We fitted the GLM using the function `glm` provided by the packages `lme4` and `lmerTest` for RStudio (Bates et al., 2014).

### 3.3. Results

#### Descriptive

The judges reported 203 instances of guilt across the guilt videos in condition 1 (respectively 161 in condition 2), with an average of three pinpoints per video (both conditions). In 16 (respectively 6 in condition 2) of those instances (8 to 3% of the total amount of guilt pinpoints identified, see Table 3.2), guilt was associated with one other emotion (see Table 3.2 for details).

**Table 3.2. Descriptive overview of guilt pinpoints identified by judges**

	Guilt with another	Guilt occurring with (details)			Guilt alone	Grand Total
		Discomfort	Embarrassmen	Surprise		
Condition1: Occurences	16	7 (43%)	7 (43%)	2 (12%)	187	203
Context Percentages	8%				92%	
Condition2: Occurences	6	2 (33%)	2 (33%)	2 (33%)	155	161
No Context Percentages	3%				96%	
Overall Occurences	22	9	9	4	342	364
Percentages	6%	2.47%	2.47%	1.10%	94%	

*N.B. percentages in bracket indicate the proportion when guilt was reported with another emotion*



Looking at the impact of context, judges were more reliable in the presence of context. Indeed, we found a positive correlation between the judged guilt and the reported felt-guilt in the context condition ( $r=0.276$ ,  $n=56$ ,  $p = 0.0435$ ) but not in the no context condition ( $r=0.261$ ,  $n=56$ ,  $p = 0.0734$ ). However, the presence of context only explained 8% of the accuracy ( $r^2 = 0.076$ ) and the absence of context explained 7% of the accuracy ( $r^2 = 0.068$ ). The current data might not hold the appropriate power to identify a strong effect of contextual information on the accuracy of the judgement.

### Principal Components Analysis

Examining the AU/ADs produced during the guilt induction task (guilt videos), judged guilt clustered with AU4 Brow Lowerer) and Face Touch in a factor explaining 9.27% of the variance (see Table 3.3 for details).

**Table 3.3. Factor loadings of the Principal Component Analysis (PCA)**

All participants							
Rotated Component Matrix							
AUs loaded	Component						
	1	2	3	4	5	6	7
AU23/24	<b>0.817</b>	0.056	-0.047	0.076	0.089	-0.069	0.031
AU17	<b>0.779</b>	0.115	-0.060	-0.020	0.110	0.323	-0.045
AU26/27	<b>-0.564</b>	0.133	-0.069	0.055	0.402	-0.110	0.072
AU14	<b>0.519</b>	-0.079	0.240	0.154	-0.087	-0.414	0.118
AU2	0.036	<b>0.910</b>	0.032	-0.117	-0.068	0.156	0.006
AU1	0.007	<b>0.900</b>	0.218	0.031	-0.022	-0.169	0.041
FaceTouch	0.023	0.088	<b>0.762</b>	-0.133	0.036	0.013	-0.134
AU4	0.007	0.136	<b>0.664</b>	0.012	0.181	0.250	-0.002
JudgedGuilt	-0.058	0.069	<b>0.517</b>	0.052	-0.509	-0.160	0.371
AU64	0.039	-0.057	-0.111	<b>0.782</b>	-0.055	0.186	-0.042
AU5	-0.034	0.020	-0.008	<b>-0.756</b>	-0.087	0.144	-0.016
AU1011	0.031	-0.106	0.241	0.045	<b>0.813</b>	-0.098	0.120
AU54	0.153	-0.034	0.251	0.062	-0.117	<b>0.778</b>	0.084
NeckTouch	0.006	0.033	-0.099	-0.033	0.080	0.062	<b>0.934</b>

*Extraction Method: Principal Component Analysis; Rotation Method: Varimax with Kaiser Normalization. Numbers in bold indicate the higher loading for each AU and the contribution to a specific factor*

### Impact of contextual information

As expected, the presence of contextual information improved the judgement of guilt ( $\beta = -14.24$ ,  $SE = 2.55$ ,  $p < 0.0001$ ) but there was no effect of congruence between ethnicity of the judge and the individual in the video on judgement of guilt ( $\beta = 0.22$ ,  $SE = 0.87$ ,  $p = 0.549$ ). Finally, judges reported a higher level of perceived guilt the more individuals in the video produced AU4 and Face Touch ( $\beta = 5.76$ ,  $SE = 2.08$ ,  $p = 0.0057$ ).

### 3.4. Discussion

This study aimed to identify the impact of contextual information on the perception of guilt. We also aimed to identify which facial movements were perceived as guilt in a USA-based sample by reproducing a study conducted on a UK-based sample (Julle-Danière et al., under review; see Chapter 2). The results from the previous study were replicated: judges relied on AU4 (Brow Lowerer) and face touching to identify guilt on a face. We also revealed an important impact of contextual information on people's ability to perceive guilt. In the presence of contextual information, we found a positive relationship between the level of self-reported guilt and the extent this individual was judged as feeling guilty by others. This confirmed the importance of context in the interpretation of the facial expression of guilt. We showed that USA-based observers relied on the exact same facial signals to identify guilt on a face: judges associated the experience of guilt with the production of a frown (AU4 Brow Lowerer) and face touching (Ekman & Friesen, 1978). Finally, we identified a major effect of context on the perception of guilt: in the absence of contextual information, participants made poorer judgment and were not able to perceive guilt as well as when provided with the context surrounding the videos (i.e., in which situation was the stimulus individual in the video).

In real life, facial expressions always occur in context, which is why in our first study we introduced the stimuli with the general context the individuals were in (Julle-Danière et al., under review); however, previous judgement studies have demonstrated that people can reliably identify some facial expressions of primary emotion in the absence of context (Crivelli et al., 2017; Jack et al., 2016; Keltner & Buswell, 1996). In those studies, most judges agreed on the emotional state expressed on the face (reliable judgement) and selected the targeted emotion accurately (valid judgement) (Crivelli et al., 2017; Jack et al., 2016). Introducing contextual information (i.e., written information, body posture, social environment) to judgment of primary emotions revealed a conflicted picture, revealing the impact of incongruent context on the understanding of the facial expressions presented (Aviezer et al., 2008; J. M. Carroll & Russell, 1996; Fernandez-Dols et al., 1993; Hess et al., 2016; Matsumoto & Sung Hwang, 2010). In those experiments, judges made a reliable inaccurate judgement, i.e. judges agreed on what emotion was expressed but they selected a non-target emotion (Aviezer et al., 2008; Hess et al., 2016). Less research has been conducted on facial expressions of secondary emotions (Keltner & Buswell, 1996), and when doing so they usually presented the stimuli without context. Here we found that context influenced greatly whether people were able to identify a facial display of guilt, contributing to the idea that secondary emotions are more idiosyncratic and context-dependent than primary emotions.

Together, our results suggest that the perception of guilt is consistent over a large sample of individuals, living in different continents but presenting high demographic similarities (both native and non-native higher education students). It confirms that guilt seems to be expressed reliably and consistently on the face. Our results do not allow us to claim for universality of facial expression of guilt, but they do however tend to demonstrate consistency in facial signals used to perceive guilt on a face. It could be interesting to conduct our production study (Julle-Danière et al., under review - Study 1) with a US-based sample to identify potential differences in the facial expression associated with experiencing guilt. We also showed that facial expression of

guilt is better interpreted when the situational context is provided alongside the facial expression. It supports the trend for ecologically valid judgment research (Matsumoto & Sung Hwang, 2010), as a facial display is rarely occurring without context in real life.

#### **4. Perception of dynamic facial expressions of 'guilt'**

##### **Abstract**

Secondary emotions such as guilt are not thought to be associated with a clear and distinct facial expression, but a recent study identified a pattern of nonverbal behaviour associated with the experience of guilt. However, it is unclear how people interpret this display in the absence of context. Using dynamic, ecologically valid stimuli, we show that people do not reliably identify a facial expression associated with guilt when using explicit emotion or action tendency labels but seem to distinguish the expression of guilt from other emotional displays when using dimensional ratings (aroused-asleep; pleasant-unpleasant). Therefore, secondary emotions may be reliably interpreted by observers, but not necessarily using categorical, discrete labels. This tells us secondary emotion might still be associated with reliable signals, but that these are not necessarily conceptualised as discrete phenomena. Contextual information might be necessary to improve the understanding of facial displays of secondary emotions.

**Keywords.** Facial expression, guilt, emotion, Basic Emotion Theory, Behavioural Ecology View

#### 4.1. Background

Guilt is an emotional and cognitive experience, arising when someone feels that s/he did something wrong. It is classified as a moral, self-conscious emotion (Haidt, 2003), and is one of the most social, other-oriented emotions that people can experience throughout life (J. Carroll, 1985; Cryder et al., 2012; Tangney, 1999). Guilt is often mistaken or mislabelled as shame, or sometimes embarrassment and research has tried to differentiate between those, not only in terms of psychological experience but also in the behavioural signal (Keltner & Buswell, 1996; Tangney, 1999; Teroni & Deonna, 2008). Behavioural responses to embarrassment and shame have been clearly identified (Izard, 1977; Keltner, 1995; Lewis et al., 1992): embarrassment display is marked by gaze down, controlled smiles, gaze shifts, and face touches (Keltner, 1995), whereas a shameful display is marked with head and gaze down (Izard, 1977; Keltner, 1995; Lewis et al., 1992), but a similarly discernible display has not been associated with guilt. Embarrassment serves a reconciliatory and appeasement function, reconciling social relations following transgressions of social norms, and shame arises following failure to live up to expectations associated with the "core-self" (see Keltner & Anderson, 2000 for review; Keltner & Buswell, 1996). Embarrassment arises following social transgressions of social conventions, a set of culturally specific rules governing social interactions (e.g., manners, outfit). Shame follows violations reflecting the individual's character, affecting one's social image (e.g., hurting someone's feelings, being a bad person; Keltner, Young, & Buswell, 1997; Keltner, 1995). Guilt also seems to be distinguishable from these emotions in terms of its function and experience. The antecedents leading to guilt have been identified as violation of moral rules governing behaviours towards others (i.e., lying, cheating, stealing; Keltner & Anderson, 2000), and when the person feels bad about the transgression itself (Cryder et al, 2012). The main outcomes resulting from guilt seem to be prompting attempt to make amend and correct past transgressions (Cryder et al, 2012; Tangney, 1999). Therefore, it is possible that guilt also has a recognisable display. A recent

study suggested that there is a facial expression associated with guilty feelings, which correlated with judgements of guilt by others (e.g. How guilty is this person?, Julle-Danière et al., under review), but the cognitive processes observers are using to interpret this display, and how the relates to meaning, are still to be investigated.

Whether emotions (and which emotions) are associated with universally produced and recognised facial expressions is debated (Crivelli & Fridlund, 2018; Keltner, Tracy, Sauter, & Cowen, 2019) and relate directly to different theories about emotion. The classic and largely dominant view is that primary, basic emotions (happiness, sadness, anger, surprise, disgust and fear - Ekman & Cordaro, 2011) are considered innate to all human populations and universally expressed and understood (Brown, 1991; Ekman & Friesen, 1969), and so likely resulting from natural selection (Ekman & Keltner, 1970). Within the same view, secondary emotions (of which guilt is one, along with embarrassment, shame, and contempt) are thought to differ significantly between cultures (Bedford & Hwang, 2003; Matsumoto et al., 2005), their expressions subject to specific cultural display rules (Ekman & Friesen, 1969; Matsumoto et al., 2005), and acquired and developed gradually during childhood (Tangney, 1999). In sum, secondary emotions are thought to be more idiosyncratic and context-dependent in terms of both production and perception.

Different cultures could rely on specific facial clues, which are functional in a given culture and for a given emotion, but do not necessarily translate to a different culture (Jack, Garrod, et al., 2012; Keller & Otto, 2009), and whether people experience a specific emotion seems to be related to their cultural norms (Mesquita & Leu, 2007). Emotion could be experienced more in one culture (i.e., anger in European countries) and people in different cultures can experience emotions differently (i.e., anger associated with a feeling of control in European countries but a feeling of guilt in East-Asian countries), and these processes do not seem genetically fixed. Indeed, previous research has shown a stronger impact of culture

and environment (i.e. nurture) on the production of facial expression than the impact of ethnicity (i.e. nature; Camras et al., 2006). Chinese girls adopted by Western families displayed a level of expressivity significantly higher than Mainland Chinese girls and Chinese American girls (both living in Chinese families), revealing an important influence of culture and family attitudes on facial expressivity. Possible innate differences between ethnicity are superseded in favour of cultural influences (Camras et al, 2006). Moreover, recent work has shown an ability for people changing cultural environment to 'adapt' to their new culture by experiencing emotional acculturation and they can begin to experience emotion accordingly to their new cultural standards (De Leersnyder, 2017). The cultural impact on facial expressivity is thus not limited to early childhood exposure, but also affects the expressiveness of adults changing culture. Some research has been conducted on the acculturation in emotional expression (for review see: Cordaro et al., 2018), i.e. whether the expression or perception of a given emotion changes over time. Previous research demonstrated that individuals from different cultures used different cues to identify emotional facial expressions (Jack, Caldara, et al., 2012; Yan et al., 2017) or produce different facial expressions in the same situation (De Leersnyder & Mesquita, 2015). However, it has also been shown that time spent in a culture improved the recognition of a culturally specific facial expression (Elfenbein & Ambady, 2003), suggesting that people can become acculturated to different styles of emotional recognition.

The evidence for strictly universal primary emotions and culturally variable secondary emotions has been challenged and scientist still disagree on what exactly the face communicates even when there is a uniform expression: basic emotions, action tendencies, or something entirely different (Crivelli & Fridlund, 2018; see Crivelli & Fridlund, 2019 for review; Hinde, 1985; Nelson & Russell, 2013; Parkinson, 2005). Over the last 50 years, a body of research has demonstrated cultural differences in the expression (Chen & Jack, 2017) and perception of facial expression of primary emotions (Elfenbein et al., 2007; Jack, Caldara, et



al., 2012), often comparing European and Asian cultures. The universality in the perception of basic emotion has also been found to be method-bound (Fridlund, 2017; Gendron, 2017). Specifically, using forced-choice categories seems to amplify 'correct' responses and mask the noise with which expressions are spontaneously received (Gendron, 2017; Gendron, Crivelli, & Barrett, 2018). Recent studies (Crivelli et al., 2016, 2017) tried to address this issue by using "prototypical" expressions of emotion in a design comparing the Basic Emotion Theory (BET) classical approach (Ekman & Friesen, 1969) to the Behavioural Ecology View (BECV; Fridlund, 2017).

The Basic Emotion Theory (BET - Ekman, 1992a, 1992b; Keltner, Tracy, et al., 2019) argues that facial expressions are signals indicating the emotion or internal state of the sender (see Keltner, Tracy, et al., 2019 for review). As such, facial expressions are a genuine expression of the sender's internal state (regardless of how this relates to social function). The Behavioural Ecological View of facial expressions (BECV - Crivelli & Fridlund, 2018; Fridlund, 1994, 2017) argues instead that facial expressions indicate the sender's most likely future behaviours, which the authors argue is better aligned with evolutionary theory. As signals of action, facial expressions can benefit both the sender and receiver by reducing the need for conflict when interests are declared openly (Baumeister et al., 1994; Ekman & Friesen, 1971; Fridlund, 1994, 2017; Tomasello, 2008; Waller et al., 2016); as such, both primary and secondary emotions could theoretically be associated with specific, readable and recognisable, facial signals.

Recent sets of studies brought evidence that emotional facial displays could be reliably understood using social motives and action tendencies in both Western societies and small scales communities (Crivelli et al., 2016, 2017; see Gendron et al., 2018 for review; Shuman, Clark-Polner, Meuleman, Sander, & Scherer, 2017). Using still photographs of facial expressions in a forced-choice paradigm, Crivelli et al (2016) found that two of the displays

were reliably understood in terms of social motives: social invitation (which could be associated with happiness) and rejection (“disgust”). They also highlighted that some facial displays were more reliably identified using social motives compared to emotion words (nose-scrunching face was preferably attributed to rejection than disgust; Crivelli et al, 2016). In the second set of studies, Crivelli et al (2017) compared different categorisation method, using photographs of spontaneous emotional displays. Comparing free label, dimensional ratings, and forced-choice categorisation, they showed that Trobrianders categorised more reliably the facial displays when using dimensional ratings, i.e., defining the emotional states based on valence and arousal (Crivelli et al, 2017). Shuman and colleagues (2015) used posed dynamic stimuli, presented with no contextual information, in a set of categorisation tasks and found that action tendency labelling was the least accurate methodology. They hypothesised that it could be due to either the labels used or the absence of contextual information. Taken together, those two studies suggest that bi-dimensional scales assessing core affect might play a role in the interpretation of the face, highlighting what is universal in facial expressions (Crivelli et al., 2017).

### **Present investigation**

Our study aimed to build onto those studies by using spontaneous dynamic displays of emotions. We investigated how ecologically valid stimuli were perceived by observers from different cultural backgrounds and tested which categorisation methods (emotion words - BET-, action tendencies -BECV-, or quantitative ratings on multiple scales; Fridlund, 1994) led to the highest recognition. We used displays of guilt compared to displays of basic emotions. Instead of still photos, we used short videos presenting a dynamic facial expression of a given emotion, reflecting real-life experiences. Moreover, we did not provide any contextual information with the dynamic displays to test whether the face alone conveyed the predicted message. We also examined cultural differences as previous studies have revealed marked cultural differences in accuracy (Gendron et al., 2018; Jack, Caldara, et al., 2012; Mesquita &

Frijda, 1992; Russell, 1994). We hypothesised that spontaneous dynamic facial displays of secondary emotions would be more miscategorised than dynamic displays of posed primary emotions. We also hypothesised that displays of primary emotions would be understood in terms of both emotion words and social motives. Finally, we hypothesised that guilt could be best described or identified using the arousal and pleasantness dimensions.

## 4.2. Methods

### Observers

Two hundred and forty-three participants (135 female, 105 male, 3 preferred not to say) were recruited online using the online survey recruitment system Prolific.ac (Palan & Schitter, 2018). Participants' ages ranged between 18 and 74 years, with a median of 29 and an average of 33.4 years ( $SD = 12.56$  years). Participants were recruited from both Western and Asian countries, with 112 participants self-identifying as of Asian ethnicity (Asian, Chinese, Pakistani, Indian, or other Asian backgrounds), 124 participants self-identifying as of White ethnicity (White British/Irish/Scottish/other), and seven participants self-identifying as any other ethnicity (Hispanic, other, or preferred not to say). All observers gave written informed consent and received monetary compensation for taking part. The projects have been reviewed and approved by the Science Faculty Ethics Committee (SFEC) from the University of Portsmouth (Appendix 4).

### Demographics and acculturation measures

Participants were asked to indicate their birth country, as well as the birth countries of their parents and grandparents. Participants who reported being born in an Asian country were asked to fill out a four-item acculturation scale measuring the degree of *social contact*

*with members of the host culture* (see Appendix 1, Asian questionnaire; De Leersnyder, Mesquita, & Kim, 2011). Participants reporting being born somewhere other than Asia were asked to fill out a four-item questionnaire measuring the degree of exposure to and engagement with non-Western cultures (see Appendix 1, Western questionnaire; De Leersnyder et al., 2011; Delis et al., 2016). We excluded eight participants who were born outside of Asia and did not identify as Asian but had spent over three months in Asian countries, leaving us with 124 Euro-American participants. We split participants identifying as Asian into two groups: Asian participants born and raised in Asia (46 Asian participants) and Asian participants born in Western countries but of Asian descent (65 Euro-Asian participants). All analyses were conducted on these 235 participants.

## **Stimuli**

The videos used in this study were collected in Chapter 2 – Study 1. We selected 36 videos displaying seven FACS-coded facial expressions (six basic emotions: “angry”, “disgust”, “fear”, “happy”, “sad”, “surprise”, and one secondary emotion: “guilt”). Facial expressions of basic emotions were collected in Chapter 2 – Study 1 (Figure 2.1: 2). Non-verbal displays of guilt were collected during a social interaction acting as an induction task (Figure 2.2:4). Among the 36 selected videos, 12 presented basic emotion (two videos of each basic emotion, one White and one Asian participant), and 24 presented guilt displays. Cultures were equally distributed for each expression (with the exception of fear, see below for details).

**Selection of the posed basic emotion videos.** The original basic emotion videos (211) collected in Chapter 2 – Study 1 (Figure 2.1:2) were used in a screening study where we asked eight participants to watch between 90 and 211 videos, at home, in their own time, and taking as many breaks as they wanted to. For each video, we asked participants to indicate

which emotion (anger, sadness, fear, disgust, happiness, or surprise) they thought the person in the video was displaying and whether the video was a good/clear example for that emotion. Each video was judged by five participants. We then attributed an accuracy score to the judges for each video (1 if they selected the predicted emotion label, 0 otherwise) and calculated the hit rate per video (number of accurate judges/total number of judges that viewed the video, i.e. five). Then, we looked at videos with a hit rate of 100%. At this stage, none of the videos displaying an Asian person posing fear met the criteria; the highest hit rate found for an Asian-fear video was 60%, which we thought was too low to be considered a good example. We selected two videos for each emotion, selecting those that most judges qualified as a good example of the displayed emotion. We selected five stimuli presenting Asian videos (anger, disgust, happiness, surprise, sadness) and seven stimuli displaying European videos (anger, disgust, happiness, surprise, sadness, 2xfear).

**Selection of the spontaneous guilt videos.** To select good examples of displays of guilt, we used a similar screening procedure to the one described in Chapter 2 (Study 2; Jullé-Danière et al., under review). We recruited 114 participants; each viewed 20 videos (16 guilt videos and four control/neutral videos; see Chapter 2 – Study 2 for details). After each video participants were asked to indicate how much of each of five presented emotions (embarrassment, discomfort, guilt, surprise, or other) they thought the person in the video was feeling. We isolated 41 best examples of displays of guilt as videos of people reporting the highest level of guilt (felt guilt) and that received the highest rating of guilt (judged guilt). We combined those ratings to facial movements produced (FACS coding) and associated with guilty feelings to narrow down our selection to 12 good examples. We repeated the same selection procedure with videos displaying people that reported a medium to low guilt or that were judged as experiencing an average level of guilt.

In total, we selected the 12 emotion videos and 12 guilt videos that received the strongest ratings of emotion from the largest number of participants. We also selected 12 guilt videos presenting mixed results (ambiguous examples) to have control videos presenting spontaneous facial expressions differing from the displays of guilt. We cropped all videos to 4 seconds before and after the apex of the expressions displayed (all videos were 8 seconds long).

**Table 4.1. Emotion words and action tendencies used in the forced-choice tasks**

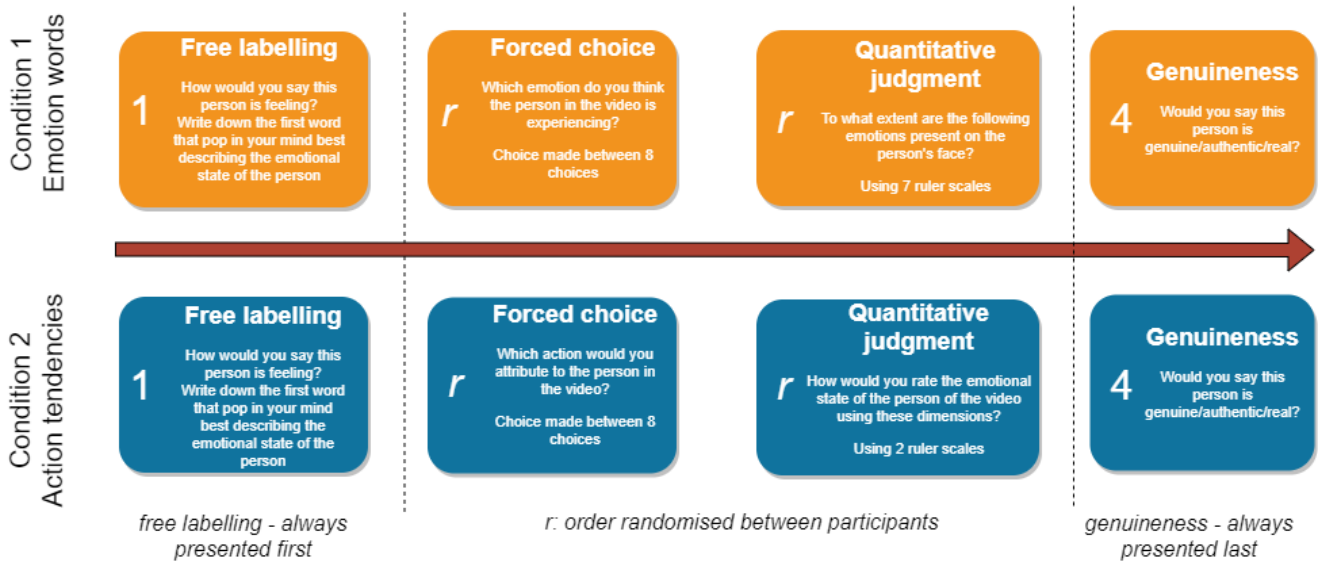
<b>Emotion words</b>	<b>Action tendencies</b>
Happiness	Ready to greet a friend
Sadness	Waiting for someone to be there to help and support her/him
Anger	About to attack someone or to yell
Fear	Ready to flee
Disgust	Trying to prevent/stop sensory contact
Surprise	Urgently attending to the current situation
Guilt	Ready to make up for what was done

*Action tendencies labels created based on Crivelli et al. (2016); Fontaine and Scherer (2013); and Yik (1999)*

## Procedure

Participants were randomly assigned to one of two conditions: the emotion condition or action tendency condition (see Table 1). After watching each video, observers performed three categorisation tasks: 1. free labelling task; 2. eight-alternative forced-choice categorisation task (all seven emotion words, or action tendency equivalents, plus “none of the above”; see Table 4.1); 3. quantitative judgment task. In the final quantitative judgment, in the emotion condition, participants had to indicate how much of each emotion they thought the person in the video was feeling, using sliding-scales (from 0-100%). In the action tendencies condition, participants had to indicate how pleasant (*Giving a sense of happy satisfaction or enjoyment; friendly and considerate; likeable; dimension: pleasant-unpleasant or positive-negative*) and arousing (*Evoke or awaken a feeling, emotion, or response; excite*

or provoke; awoken from sleep; dimension: sleep-tension OR attention-rejection) was the emotional state of the person. Order of presentation was randomised for each participant between the forced-choice categorisation and the quantitative judgment task. Finally, participants were asked to rate how genuine/authentic/real they thought the expression was, using a 5-point Likert scale ranging from “definitely not” to “definitely yes” (see Figure 4.1).



**Figure 4.1. General Procedure.**

A flowchart representing the different conditions of the experiment.

## Statistical Analysis

All the data analyses were performed with R (R Core Team, 2014), using functions from different packages. Generalised linear mixed models were computed and fitted with the *lmer* provided by the packages “lme4” and “lmerTest” (Bates et al., 2014). Multiple comparisons were computed with the function *glht* of the “multcomp” package (Hothorn et al., 2017). The pairwise Bonferroni comparisons were computed with the function *pairwise.t.test* from the “stats” package (R Core Team, 2014). Finally, the cross tables were computed with the *CrossTab* function of the “gmodels” package (Warnes, Bolker, Lumley, Warnes, & Imports, 2018). All GLMMs were fitted with a Gaussian error structure and estimated using REML. As

our dataset considered ratings of emotion videos by multiple judges, we included the identity of the judge and the stimulus individual as random factors in all the models to control for pseudoreplication (Hurlbert, 1984; Waller et al., 2013). The  $r^2$  for all linear models were calculated using the function *r.squaredGLMM* of the "MuMIn" package (Barton & Barton, 2019). We ran posthoc Tukey comparisons on significant interaction(s) and effect(s). We ran GLMMs to understand the variables influencing 1) accuracy and 2) quantitative judgements. We also ran another GLMM to investigate how judges perceived the genuineness of the facial expression, whether they could dissociate the posed facial expression from the spontaneous ones, and whether this affected the accuracy scores. Additionally, we computed a chi-square and a cross table to investigate the distribution of choices over the 8 facial displays (6 basic emotion, guilt, and control) presented and understand the choices made for each of the facial display.

### 4.3. Results

#### A. Accuracy: how good were judges at labelling guilt?

Data was clustered by emotion categories. We ran a GLMM to examine which variables influenced the accuracy scores. Mean categorisation accuracy of the judge by emotion was set as our response variable; the culture of the judge (3 x culture of judges), the culture of the stimulus' face (2 x culture of face), the condition the judge was in (emotion vs. action tendency), the method of categorisation (free vs. forced), and the facial expression (8 x facial expressions) were set as predictors.

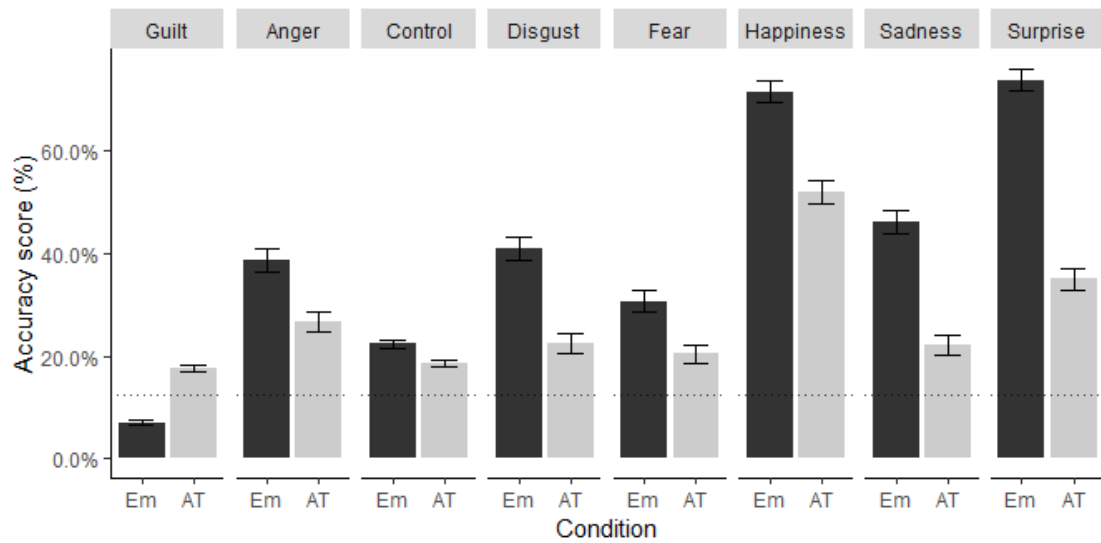
**1. Main effect – accuracy scores.** There was a significant main effect of method of categorisation on the accuracy scores: judged performed significantly better when presented with the forced-choice categorisation task than when asked to freely label the emotional



state ( $\beta = 1.51e-01$ ; SE = 1.16e-06;  $p < 0.0001$ ) and when presented with forced choice than when asked to label freely the facial expression ( $\beta = 7.97e-02$ ; SE = 1.37e-02;  $p < 0.0001$ ). We also found a main effect of facial expressions: guilt was significantly less accurately label than all other facial expressions (all  $\beta$ s  $< 3.34e-01$ ; SEs = 5.48e-02; ps  $< 0.0001$ ), including control displays ( $\beta = 8.74e-02$ ; SE = 2.93e-02;  $p = 0.00485$ ). The  $r^2$  for the full model was  $r^2 = 0.24$ . The model also showed significant interactions of the culture of judge x facial expression, condition x facial expression, and method x facial expressions. We ran posthoc Tukey comparisons to explore those interactions further.

**2. Post-hoc comparisons – accuracy scores.** Judges made significantly more errors in the action tendency condition than in the emotion condition for all facial expressions of basic emotions (all ps  $< 0.01$ ); no difference was found for control displays ( $p = 0.13$ ); for guilt displays, judges made less error in the action tendency condition than in the emotion condition ( $p_{\text{guilt}} < 0.01$ ; see Figure 4.2 and Table 4.2). Judges were significantly more accurate when categorising “surprise” using a free labelling methodology ( $p < 0.01$ ) than when using forced-choice categorisation method; moreover, judges were more accurate in categorising “anger” ( $p = 0.0318$ ), “control” ( $p < 0.01$ ), and “guilt” ( $p < 0.01$ ) using a forced-choice methodology compare to a free labelling task (see Figure 4.3 and Table 4.3). Finally, European judges were less accurate when categorising “surprise” ( $p = 0.03$ ) than Euro-Asian judges (see Figure 4.4 and Table 4.4).

**3. Genuineness.** Genuineness score was set as our response variable; the facial expression and the accuracy score were set as the predictors. All facial expressions displaying a “basic” emotion were rated as less genuine than the displays of guilt (all  $\beta$ s  $< 5.733e-01$ ; SEs  $< 1.371e-02$ ; ps  $< 0.0001$ ). Moreover, there was a significant effect of scores: the more genuine the expressed emotion was rated as, the worst the accuracy was ( $\beta = -8.885e-02$ ; SE = 2.155e-02;  $p < 0.0001$ ).

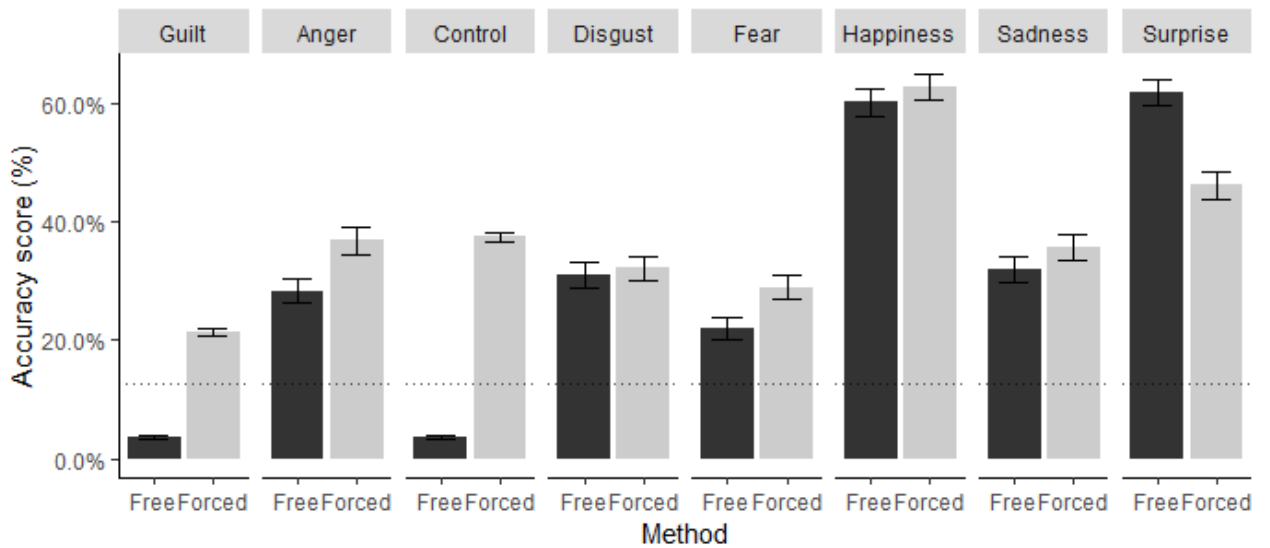


**Figure 4.2. Accuracy scores presented for each facial expression in each condition.**

The dashed line represent chance level (12.5%); Em: emotion words; AT: action tendencies. The error bars represent the Standard Error.

**Table 4.2. Proportion of judges who matched the predicted expression to the corresponding emotion or action tendency label**

Facial expressions	Prediction				P(1)-P(2)	p
	Emotion		Action tendency			
	mean	SD	mean	SD		
Anger	0.39	0.49	0.27	0.44	0.12	<0.01
Disgust	0.41	0.49	0.23	0.42	0.19	<0.01
Fear	0.31	0.46	0.20	0.40	0.10	<0.01
Happiness	0.72	0.45	0.52	0.50	0.20	<0.01
Sad	0.46	0.50	0.22	0.42	0.24	<0.01
Surprise	0.74	0.44	0.35	0.48	0.39	<0.01
Guilt	0.07	0.26	0.18	0.38	-0.11	<0.01
Control	0.22	0.42	0.19	0.39	0.04	0.13



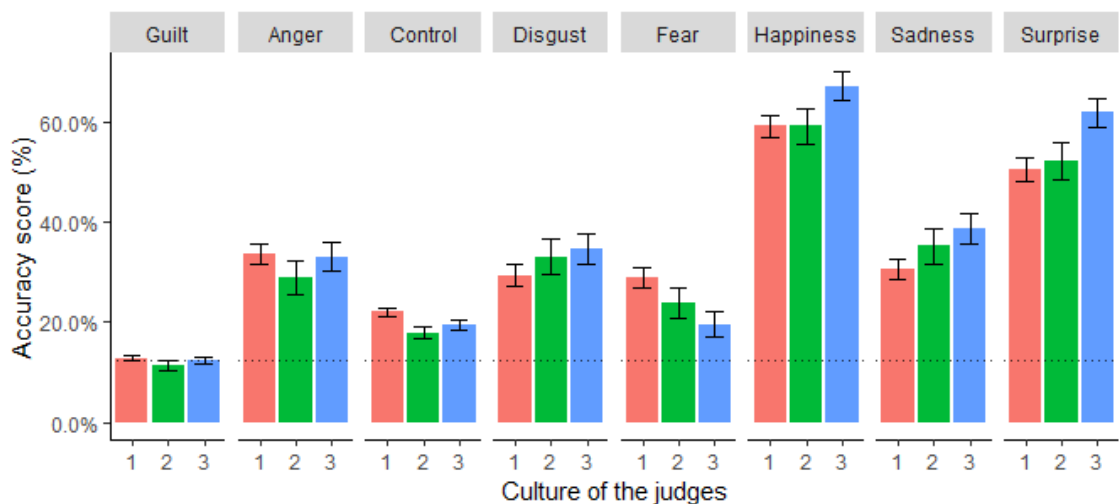
**Figure 4.3. Accuracy scores presented for each facial expression for each method used.** The dashed line represent chance level (12.5%); Free: free labelling; Forced: forced choice categorisation. The error bars represent the Standard Error.

**Table 4.3. Proportion of judges who matched the predicted expression to the corresponding emotion or action tendency label based on the categorisation method used**

Facial expressions	Prediction				P(1)-P(2)	<i>p</i>
	Free labelling		Forced choice			
	mean	SD	mean	SD		
Anger	0.28	0.45	0.37	0.48	-0.09	<b>0.0318</b>
Disgust	0.31	0.46	0.32	0.47	-0.01	1
Fear	0.22	0.41	0.29	0.45	-0.07	0.1752
Happiness	0.60	0.49	0.63	0.48	-0.03	0.9992
Sad	0.32	0.47	0.36	0.48	-0.04	0.9533
Surprise	0.62	0.49	0.46	0.50	0.16	<b>&lt;0.01</b>
Guilt	0.04	0.18	0.21	0.41	-0.18	<b>&lt;0.01</b>
Control	0.04	0.18	0.37	0.48	-0.34	<b>&lt;0.01</b>

## B. Pattern of categorisation: what did judges label guilty faces as?

**1. Overall results.** To investigate more in-depth how judges labelled each facial display, we looked at the choices made by judges. We wanted to know for each facial display which label was preferentially selected by judges. We have seen above (see Accuracy section) that judges had a poor recognition accuracy when it came to guilty displays; this set of analysis helped us understand why and how guilty displays were labelled (if judges did not select preferentially "guilt", what did they select?). Overall, judges accurately labelled facial expressions displaying one of the six "basic" emotions but made more mixed attributions for the facial expressions associated with the experience of guilt (see Tables 4.2 and 4.4 for details). We examined in details how judges from different cultural background perceived the facial displays and whether emotion words led to a better labelling of facial displays than action tendencies.



**Figure 4.4. Accuracy scores presented for each facial expression for each cultural group.** The dashed line represent chance level (12.5%); 1 – European judges; 2 – Asian judges; 3 – Euro-Asian judges. The error bars represent the Standard Error.

**Table 4.4. Proportion of judges who matched the predicted expression to the corresponding label (emotion or action tendency) depending on the culture of the judge**

Facial expressions	Prediction			P(1)-P(2)	<i>p</i>	P(1)-P(3)	<i>p</i>	P(2)-P(3)	<i>p</i>
	European judges (N=124)	Asian judges (N=46)	Euro-Asian judges (N=65)						
Anger	0.51	0.40	0.52	0.11	1.00	-0.01	1.00	-0.11	1.00
Disgust	0.46	0.45	0.54	0.01	1.00	-0.08	1.00	-0.09	1.00
Fear	0.42	0.31	0.34	0.11	1.00	0.08	0.20	-0.03	1.00
Happiness	0.83	0.82	0.86	0.01	1.00	-0.03	0.72	-0.04	1.00
Sadness	0.43	0.47	0.54	-0.04	1.00	-0.11	0.62	-0.07	1.00
Surprise	0.68	0.68	0.78	0.00	1.00	-0.10	<b>0.03</b>	-0.10	0.89
Guilt	0.11	0.11	0.07	0.00	1.00	0.04	1.00	0.04	1.00

*N.B. control displays did not have any associated predicted labels and are not presented in this table*

**2. Culture of judges.** At least 17% of judges selected “None of the above” for each facial display. When a choice was made, we found that for all judges, six out of seven facial displays were modal for the predicted emotion type (anger, fear, disgust, happiness, sadness, and surprise). The proportion of judges that selected the predicted label ranged from extremely high (happiness<sub>EU\_EA</sub> = 72%), through high (happiness<sub>EU</sub> = 62%; happiness<sub>EA</sub> = 60%), moderate (surprise > 41%; sadness > 33%; fear<sub>EU</sub> = 33%; anger > 31%; disgust > 30%), and low (fear<sub>EU\_EA</sub> = 26%; fear<sub>EA</sub> = 25%). When looking at guilty faces, the labels associated with “guilt” were not selected above chance level and judges selected mainly “Fear” (15% of Asian judges and 12% of Euro-Asian judges) and “Sadness” (15% of European judges). Those results are presented in Table 4.5.

Table 4.5. Proportion of judges matching a facial expression to a label depending on the culture of the judge

Culture Judges	Facial expressions							
	Anger	Control	Disgust	Fear	Guilt	Happiness	Sadness	Surprise
European judges (N=124)								
Anger	<b>38.31**</b>	1.88	4.03	4.03	6.12	3.23	4.03	2.82
Disgust	6.45	4.10	<b>30.24**</b>	9.68	8.47	0.81	12.50	2.42
Fear	13.31	7.59	6.85	<b>33.47**</b>	12.57	2.82	8.06	2.42
Guilt	8.87	6.59	3.63	8.06	7.73	2.42	14.92	2.02
Happiness	0.81	9.01	0.81	0.81	3.70	<b>61.69**</b>	0.81	9.27
Sadness	4.84	11.96	12.10	12.90	<b>14.65**</b>	1.21	<b>33.47**</b>	2.02
Surprise	2.42	<b>21.71**</b>	8.06	10.89	18.15	2.02	3.23	<b>44.76**</b>
None of the above	25.00	<b>37.16**</b>	34.27	20.16	28.63	25.81	22.98	34.27
Asian judges (N=46)								
Anger	<b>31.52**</b>	1.99	2.17	7.61	6.52	2.17	8.70	3.26
Disgust	14.13	3.62	<b>33.70**</b>	15.22	9.42	0.00	11.96	3.26
Fear	11.96	8.51	10.87	<b>25**</b>	<b>14.67**</b>	1.09	8.70	4.35
Guilt	15.22	7.97	4.35	8.70	8.33	1.09	9.78	1.09
Happiness	0.00	9.78	0.00	0.00	3.99	<b>59.78**</b>	0.00	5.43
Sadness	3.26	10.87	13.04	15.22	13.41	7.61	<b>35.87**</b>	2.17
Surprise	2.17	<b>24.46**</b>	10.87	8.70	20.29	1.09	1.09	<b>41.30**</b>
None of the above	21.74	<b>32.79**</b>	25.00	19.57	23.37	27.17	23.91	39.13
Euro-Asian judges (N=65)								
Anger	<b>37.69**</b>	1.03	1.54	3.85	5.77	3.08	8.46	0.00
Disgust	11.54	3.33	<b>34.62**</b>	9.23	9.62	3.08	6.15	1.54
Fear	7.69	4.10	5.38	<b>26.15**</b>	<b>12.18**</b>	0.77	3.08	0.00
Guilt	6.15	5.00	0.77	6.92	5.00	0.77	11.54	0.00
Happiness	0.77	7.69	0.77	0.00	2.05	<b>71.54**</b>	0.00	11.54
Sadness	4.62	10.64	10.00	12.31	10.38	1.54	<b>40**</b>	1.54
Surprise	4.62	<b>27.05**</b>	11.54	17.69	<b>24.23</b>	2.31	4.62	<b>52.31**</b>
None of the above	26.92	<b>41.15**</b>	35.38	23.85	30.77	16.92	26.15	33.08

Results based on cross table analyses: \*\*  $p$  significant at 0.001; \*  $p$  significant at 0.01  
 results in bold highlight the main choice made for a given facial expression; results in italic indicate the next most selected label when "None of the above" was mainly selected

**3. Emotion words: how do judges label guilt displays using emotion words?**

9% to 33% of judges selected "None of the above" for a given facial expression using emotion words. When a choice was made, we found that when using emotion words, six out of seven facial displays were correctly categorised (anger, fear, disgust, happiness, sadness, and surprise). The proportion of judges that selected the predicted label ranged from extremely high (surprise = 83%; happiness = 77%) to moderate (sadness = 49%; disgust = 44%; anger = 43%; fear = 35%). When looking at guilty faces, the emotion label associated with "Guilt" was not selected above chance level and judges selected mainly "Fear" (18% of judges; see Table 4.6).

**4. Action tendency labels how do judges label guilt displays using action tendencies?**

Overall, four of the action tendency labels were accurately associated with the predicted facial expressions [anger, fear, happiness, and sadness; N.B.: when categorising fear, the same proportion of judges selected "None of the above" and "Fear"], but 25 to 60% of the judges selected "None of the above" in this condition. The proportion of judges that selected the predicted label ranged from moderate (happiness = 51%), low (anger = 31%; fear = 25%; sadness = 13%; disgust = 21%), to extremely low (surprise = 11%). When looking at guilty faces, the labels associated with "Guilt" were not selected above chance level and judges selected mainly "Surprise" (22% of judges; see Table 4.6).

Table 4.6. Proportion of judges matching a facial expression to an emotion or action tendency label

Condition	Facial expressions							
	Anger	Control	Disgust	Fear	Guilt	Happiness	Sadness	Surprise
Emotion (N=115)								
Anger	<b>43.04**</b>	1.81	1.30	2.61	7.46	2.17	8.70	1.30
Disgust	10.43	3.33	<b>43.91**</b>	11.30	7.90	1.30	5.65	0.00
Fear	10.43	9.78	4.35	<b>34.78**</b>	<b>17.90**</b>	1.74	4.35	1.30
Guilt	12.61	5.51	3.91	9.57	8.19	1.74	13.48	0.00
Happiness	0.00	8.91	0.87	0.00	2.90	<b>77.39**</b>	0.43	6.09
Sadness	3.04	7.90	12.61	9.13	10.51	1.30	<b>48.70**</b>	0.00
Surprise	0.87	<b>23.19*</b>	10.00	15.65	18.26	2.61	1.74	<b>82.61**</b>
None of the above	19.57	<b>39.57**</b>	23.04	16.96	26.88	11.74	16.96	8.70
Action tendencies (N=120)								
Anger	<b>30.83**</b>	1.53	4.58	6.67	4.79	3.75	3.75	2.92
Disgust	8.33	4.24	<b>20.83**</b>	10.00	10.00	1.25	15.42	4.58
Fear	12.50	3.96	10.00	<b>25**</b>	8.06	2.08	9.17	2.92
Guilt	6.25	7.29	2.08	6.25	6.04	1.67	12.50	2.50
Happiness	1.25	8.68	0.42	0.83	3.68	<b>51.25**</b>	0.42	<b>12.08</b>
Sadness	5.83	14.72	10.83	17.08	15.83	3.75	<b>23.33**</b>	3.75
Surprise	5.00	<b>24.24**</b>	9.17	9.17	<b>22.15**</b>	1.25	4.58	11.25
None of the above	30.00	35.35	<b>42.08**</b>	<b>25**</b>	<b>29.44**</b>	35.00	30.83	<b>60**</b>

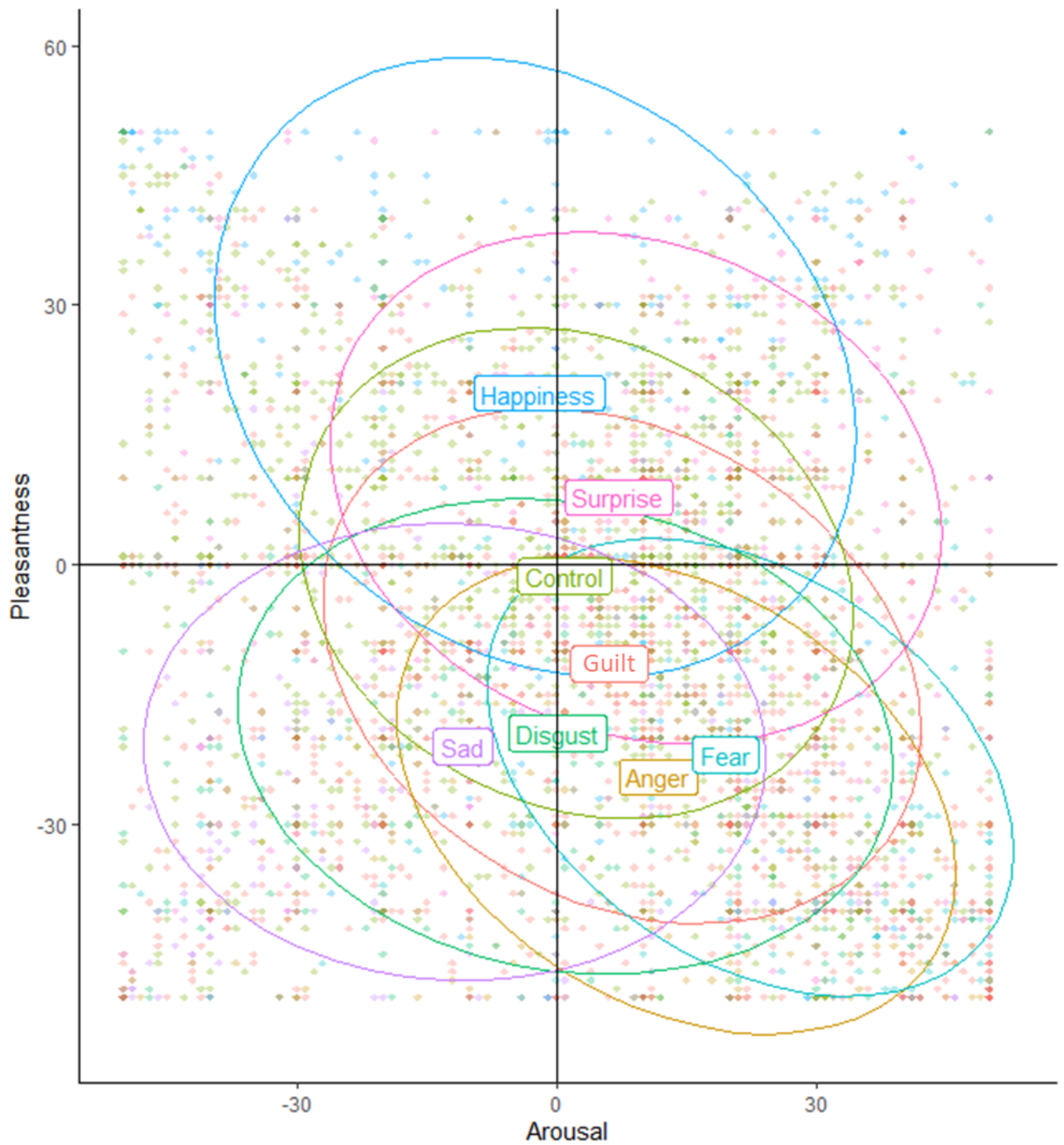
Results based on cross table analyses: \*\* *p* significant at 0.001; \* *p* significant at 0.01  
 results in bold highlight the main choice made for a given facial expression; results in italic indicate the next  
 most selected label when "None of the above" was mainly selected



### C. Quantitative judgements: how did judges categorise guilt using bi-dimensional scales or emotion rulers?

**1. Arousal and pleasantness dimensions.** We investigated the variables influencing dimensional judgements of arousal and pleasantness using two GLMMs. First, mean rating of arousal by the judge was set as our response variable, and the culture of the judge (3 x culture of judges), the culture of the stimulus' face (2 x culture of face), and the facial expression (8 x facial expressions) were set as predictors. The model was then repeated with mean pleasantness as the response variable. The  $r^2$  for the arousal model was  $r^2 = 0.26$  and the  $r^2$  for the pleasantness model was  $r^2 = 0.44$ .

The dimensional GLMMs showed a significant main effect of facial expression for both arousal and pleasantness but no main effect of the culture of the judge on either arousal or pleasantness. The GLMM looking at arousal showed that guilt displays were rated as significantly less aroused than fearful displays ( $\beta = 15.67$ ; SE = 5.89;  $p = 0.0128$ ) but as more aroused than sad displays ( $\beta = -16.97$ ; SE = 5.74;  $p = 0.0064$ ). The second GLMM looking at pleasantness showed that judges rated guilt displays as significantly less pleasant than control ( $\beta = 10.13$ ; SE = 2.87;  $p = 0.00149$ ), happy ( $\beta = 31.30$ ; SE = 5.36;  $p < 0.001$ ), and surprised ( $\beta = 19.26$ ; SE = 5.36;  $p = 0.00129$ ) displays; however, judges rated guilt displays as more pleasant than angry ( $\beta = -13.06$ ; SE = 5.36;  $p = 0.0217$ ) and fearful ( $\beta = -12.12$ ; SE = 5.50;  $p = 0.0361$ ) displays. Overall, facial expressions associated with the experience of guilt were rated with average arousal and pleasantness (see Figure 4.5 - *ratings from 0 to 100% transformed into -50/+50 intervals* - and Table 4.7).



**Figure 4.5. Scatter plot showing the ratings for each facial expression in the bi-dimensional space** Each dot represents the rating made by one judge; the ellipses regroup 67% of the distribution for each facial display; the labels with the display names (i.e. "Guilt") represent the mean of each distribution.

**Table 4.7. Mean and standard deviation of Arousal and Pleasantness ratings for each emotion type**

Facial expressions	Arousal		Pleasantness	
	Mean	SD	Mean	SD
Anger	61.78	24.06	25.47	21.35
Disgust	50.13	27.82	30.53	20.43
Fear	<b>69.48 *</b>	24.18	28.05	19.95
Happiness	47.99	26.65	<b>69.83 ***</b>	27.97
Sadness	<b>39.13 **</b>	26.37	29.00	20.23
Surprise	57.19	26.95	<b>57.79 **</b>	23.28
<i>Guilt</i>	56.10	26.06	38.53	22.66
<i>Control</i>	50.91	24.46	<b>48.66 **</b>	22.05

Comparison between rating looking at the differences between facial expression of guilt and all other facial expressions; Results based on GLMMs analyses: \* p significant at 0.05; \*\* p significant at 0.01; \*\*\* p significant at 0.001. N.B. control displays did not have any associated predicted labels and are not presented in this table

**2. Quantity of emotion.** We investigated the variables influencing the quantity rating of emotion using seven GLMMs. Mean ratings of emotion by the judge were set as the response variables, and the culture of the judge (3 x culture of judges), the culture of the stimulus' face (2 x culture of face), and the facial expression (8 x facial expressions) were set as predictors. The  $r^2$  for the full models were:  $r^2_{\text{guilt}} = 0.085$ ;  $r^2_{\text{anger}} = 0.42$ ;  $r^2_{\text{disgust}} = 0.43$ ;  $r^2_{\text{fear}} = 0.38$ ;  $r^2_{\text{happiness}} = 0.57$ ;  $r^2_{\text{sadness}} = 0.39$ ;  $r^2_{\text{surprise}} = 0.50$ .

When looking at the ratings of guilt, the GLMMs showed a main effect of culture of the judges: Euro-Asian judges judged guilt displays (post-hoc comparison:  $p = 0.0277$ ), angry displays (post-hoc comparison:  $p = 0.00152$ ), and control displays (post-hoc comparison:  $p = 0.0363$ ) as less guilty than Asian judges. Moreover, sad displays received the highest rating of guilt ( $\beta = 8.94$ ;  $SE = 3.14$ ;  $p = 0.00827$ ) and guilt displays were rated as showing more guilt than happy ( $\beta = -7.20$ ;  $SE = 3.14$ ;  $p = 0.0296$ ) and surprised ( $\beta = -8.59$ ;  $SE = 3.14$ ;  $p = 0.0107$ ) displays.

When looking at facial displays of guilt, post-hoc Tukey comparisons showed that judges rated them as showing mainly surprise (all  $p$ s < 0.01). The rating of level of guilt was lower than the rating of level of fear ( $p$  < 0.01) and sadness ( $p$  < 0.01) but higher than the level of anger ( $p$  < 0.01) and happiness ( $p$  < 0.01). All other facial displays of emotion were rated as presenting the associated emotion the most (e.g., angry videos were rated as the angriest among all videos; see Figure 4.6 and Table 4.8).

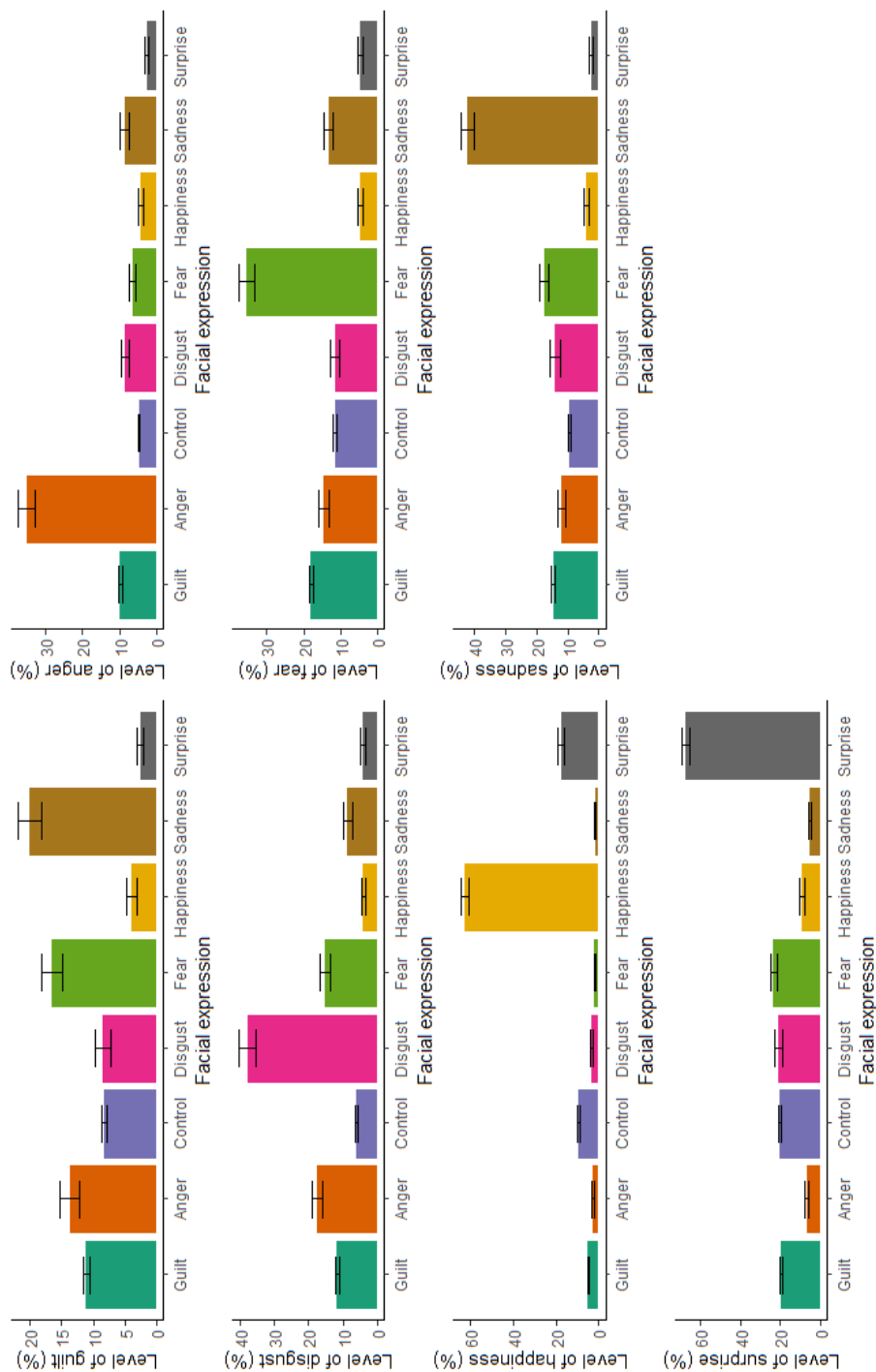


Figure 4.6. Emotion ratings for each facial expression (in percent)

The error bars represent the Standard Error.

Table 4.8. Mean and standard deviation of emotional ratings for each emotion type

Facial expressions	Anger		Disgust		Fear		Guilt		Happiness		Sadness		Surprise	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Anger	<b>35.11</b> ***	34.48	17.61	22.80	14.56	23.29	13.70	22.91	2.55	9.05	11.93	18.54	6.64	14.52
Disgust	8.52	15.48	<b>37.79</b> ***	35.70	11.43	18.27	8.44	19.80	3.51	9.88	13.95	24.99	20.78	26.63
Fear	6.42	13.38	15.26	22.32	<b>35.4</b> ***	31.30	16.47	25.80	1.84	6.55	17.60	22.85	23.19	26.25
Happiness	4.23	12.65	4.08	10.32	4.57	11.17	3.9 *	11.96	<b>62.29</b> ***	30.86	3.98	10.30	8.92	16.39
Sadness	8.61	18.45	8.60	17.90	13.22	18.78	<b>20.04</b> **	27.83	1.70	5.87	<b>42.22</b> ***	34.25	5.13	13.22
Surprise	2.56	8.61	4.25	11.48	4.75	11.59	2.51 *	7.64	17.42	24.93	2.45	7.34	<b>67.60</b> ***	30.86
<b>Guilt</b>	<b>9.72</b> **	19.05	11.67	19.51	<b>18.01</b> ***	23.52	11.10	20.29	<b>4.99</b> **	13.45	<b>14.65</b> **	22.12	<b>19.54</b> ***	25.41
Control	4.71	11.52	6.10	13.44	11.48	17.97	8.27	17.11	9.55	18.29	9.49	17.78	20.30	24.74

Numbers presented in this table are the mean rating on each scale for each facial expression. Numbers in bold represent the facial expression rated the highest on each emotion scale; numbers in red show comparison between ratings in facial displays of guilt; numbers in italics show comparisons in ratings of guilt between all facial Results based on post-hoc Tukey comparison: \*  $p$  significant at 0.05; \*\*  $p$  significant at 0.01; \*\*\*  $p$  significant at 0.001

#### 4.4. Discussion

In this study, we examined people's understanding of facial displays of guilt in the absence of context. We wanted to compare how people described guilty faces using labels (emotion vs action tendency), as well as via dimensional space ratings (arousal-asleep/positive-negative axis) and quantitative emotional ratings. We compared performance on various categorisation tasks to find out how a dynamic, spontaneously produced, guilty face would be described and understood. We also compared performance with guilt expressions to facial expressions of primary emotions in the absence of context. We looked at cultural differences to determine whether patterns were culturally variable or uniform. In sum, our results show that even in the absence of explicit contextual information, people are able to distinguish between facial expression using action tendencies as well as emotion terms. They did so reliably for faces displaying so-called "basic" emotional states (Ekman & Cordaro, 2011) but failed to do so when judging a secondary expression, namely guilt. Finally, people seemed to be able to differentiate all seven emotional states on a bi-dimensional plan, placing the experience of guilt at a mid-arousal/mid-pleasant point.

Building on previous studies looking at the decoding of ethnically varied faces by ethnically varied observers (Jack et al., 2009; Jack et al., 2016), we found a main effect of the culture of the observer on the accuracy of the judgement: observers with a mixed ethnic background (Euro-Asian judges) were more accurate than judges from a European ethnic background when it came to accurately label facial displays of surprise and sadness. However, European judges did better at labelling fearful displays than Euro-Asian judges. This could be because we only presented European faces displaying fearful expressions (due to selection criteria; see Method section for details), leading to an in-group advantage (Elfenbein & Ambady, 2003). That being said, the interaction culture of judge x culture of video was not significant in our model.

Following previous research, we found that people relied significantly more frequently on emotion attribution rather than on action tendencies for the facial expressions previously categorised as emotional displays of so-called “basic” emotion (anger, fear, disgust, happiness, surprise, and sadness; Crivelli et al., 2016; Shuman et al., 2017). Moreover, building on previous research, judges could accurately attribute both emotions and action tendencies to five of the facial expressions of emotional displays of “basic” emotion (anger, fear, disgust, happiness, and sadness), replicating Shuman and colleagues’ findings (2017). We found mixed results regarding the remaining two emotional displays studied: “Surprise” displays were accurately labelled in the condition using emotion words but not when using action tendency labels; “Guilt” displays were not accurately labelled, regardless the condition presented to the judges (emotion vs. action tendency labels).

We cannot dismiss that the results could be explained by the social meaning items used in this study. Even though we tried to base our action tendency items on labels used in previous research, in the case of “Surprise” and “Guilt”, we had to create new labels matching the grammar and structure of our other labels (see Table 4.1), without using emotion words or writing sentence using “I/me” as previously done (Crivelli et al., 2016; Yik, 1999). The present study was a way to test whether those new labels were appropriate to convey the predicted emotional state. We can conclude that five of our labels conveyed the predicted emotional state but neither the new label for “Surprise” nor the label for “Guilt” conveyed reliably the predicted social message. More specific or varied labels might have been more appropriate to describe how people perceived those guilt displays. Alternatively, in the absence of context provided, judges might not be able to understand the facial expression associated with surprised and guilty feelings as carrying a social meaning.

Our results supported previous research comparing free labelling to forced-choice methodology (Fridlund, 1994; Russell, 1994): people were more accurate in their judgements when using a forced-choice categorisation than when asked to freely report the emotional state of the stimulus. This is in line with previous studies, highlighting the increased accuracy in forced-choice categorisations (Crivelli et al., 2017; Gendron, 2017). However, this pattern was not consistent for all seven emotional states presented to the observers: when looking at facial displays associated with guilt, anger, and fear, observers did better using a forced-choice categorisation than a free labelling; but when looking at surprised displays, the accuracy was better using a free labelling methodology. We are uncertain why we found this switch of results. When analysing the data from free labelling, we took a conservative approach where only the exact term and associated noun/adjective would be considered as correct (e.g., anger/angry, happy/happiness/smile/smiling). This could explain the results for guilt, anger, and fear. Regarding surprise, we noticed that some judges reported “surprise” in the free labelling part but then selected another label in the forced choice categorisation. It could be due to the action tendency label used to describe a “surprise” situation, as discussed above. Judges failed to associate it with the predicted expressions, leading to an overall better judgement in the free labelling section than in the action tendency forced choice categorisation.

Finally, we looked at facial displays of guilt using quantitative ratings on multiple scales (Fridlund, 1994, 2017) to explore how people perceived facial expressions associated with a secondary emotion. On the one hand, judges reported that several emotions could be experienced by the person presenting a guilty display: on a scale from 0 to 100%, judges reported multiple emotions as present on guilty faces but only at low levels, with the highest rated emotion being “surprise”. When looking at other emotional facial displays, judges did attribute the highest rating to the predicted emotion (see Figure 4.6). On the other hand, judges perceived the guilty faces as medium arousal-medium pleasantness, putting the



experience of guilt right in the middle of the dimensional space (Figure 4.5). Moreover, judges rated examples of guilty faces as significantly less pleasant than control displays. We also found a trend for examples of guilt to be more aroused than control displays. Thus, we can assume the experience of guilt was not understood as a neutral face and judges considered the experience of guilt as presenting low excitement (i.e., medium to low arousal) and small variation in pleasantness (medium pleasantness). To go further, it would be interesting to consider using more than two dimensions in order to emphasise the difference in perception between guilt displays and control displays. Previous research have suggested that using a 4-dimension plan would provide better understanding as to how each facial display is perceived (Fontaine, Scherer, Roesch, & Ellsworth, 2007).

In this study, we presented judges with dynamic facial displays previously associated with one of seven emotional states. Our stimuli were not standardised in form but each stimulus lasted for 8-sec. Facial displays of guilt were genuine facial expressions associated with real-life experience of guilt. Although it reduced the level of standardisation of our stimuli, it improved the ecological-validity. It is, however, possible that the guilt displays presented were too complex or ambiguous to be understood without any contextual information. Context has been shown to mediate the understanding and social impact of social displays (Crivelli & Fridlund, 2018). Removing contextual information might have impaired the social message conveyed by guilt displays and impaired judges' ability to understand the situation when having to choose labels (Shuman et al, 2017). Moreover, the facial behaviours and AUs/ADs present in guilt displays were not unique to guilt expressions and might have hindered the recognition of this expression. Indeed, previous research has showed that AUs were not only present in various emotional displays, but were also not characteristic of a specific emotional dimension (Mehu & Scherer, 2015). It seems however that the facial expressions of guilt could convey a message different from the other facial displays presented in this study, as shown by the dimensional rating analysis. Our findings

could encourage researchers to try to map facial displays onto bi-dimensional plans, which could provide an alternative way to discriminate between facial displays of emotional states (Fridlund, 1994, 2017).

## **5. Eye fixation patterns when viewing 'guilt' facial expressions**

### **Abstract**

Being able to accurately perceive and interpret facial expressions is a key element of human interactions. Examining patterns of eye movement when viewing facial expression can help understand the processes underpinning facial expression discrimination. Here, we investigated the eye fixation patterns when viewing 'guilt' facial expressions in comparison to those exhibited when viewing 'basic' emotions. We used dynamic stimuli of spontaneous displays associated with genuine feelings of guilt. We found that 'guilt' facial expressions were examined in a similar manner to displays of the basic emotions. Observers exhibited scanning patterns that overlapped with some of the previously identified patterns used when looking at facial signals of fear. It is, therefore, possible that observers perceived emotional displays of guilt in terms of valence and arousal, rather than as a specific emotion. Finally, we found cultural differences in the viewing patterns of guilty displays. This is the first study looking at scanning strategies associated with signals of guilt.

**Keywords.** Facial expression, guilt, cultural differences, emotion, eye movements, dynamic

## 5.1. Background

The face is at the core of human relationships and complex social interactions are facilitated by the ability to perceive and accurately interpret facial expressions. During social interactions, people thus need to be able to perceive brief and dynamic successions of expressions and accurately interpret them to adjust their behaviours accordingly. However, what information is actually conveyed by the face remains debated (Barrett, Adolphs, Marsella, Martinez, & Pollak, 2019; Barrett et al., 2007; Cowen & Keltner, 2017; Ekman & Cordaro, 2011; Fridlund, 2017) and which facial movements do and do not convey information is unclear.

The classic and still dominant view - the Basic Emotion Theory (BET) - is that the face conveys emotion and that primary, basic emotions (happiness, sadness, anger, surprise, disgust and fear - Ekman & Cordaro, 2011) are innate to all human populations and universally expressed (Brown, 1991; Ekman & Friesen, 1969). In contrast, secondary emotions (of which guilt is one) are thought to differ significantly between cultures in their appraisals (Bedford & Hwang, 2003; Matsumoto et al., 2005), their expressions subject to specific cultural display rules (Ekman & Friesen, 1969; Matsumoto et al., 2005), and acquired and developed gradually during childhood (Tangney, 1999). Secondary emotions are more idiosyncratic and context-dependent, which is why it has been harder to identify specific facial movements associated with the experience of those emotional states. The BET supports the idea that emotion can be separated into specific categories (Ekman & Cordaro, 2011), each category differing in their appraisal and behavioural outcomes (e.g., facial expressions). As a result of this, researchers have been looking for emotional "signatures" or "fingerprints", i.e., physical (facial expressions) and physiological changes that reliably indicate the presence of one given emotion (Barrett et al., 2007; Gross & Barrett, 2011; for review, see Siegel et al., 2018). An alternative view relies on the interpretation of facial expressions as responses to a type of situation or as part of an instrumental action

(Crivelli & Fridlund, 2019; Fridlund, 1994, 2017). This view distanced itself from the Darwinian definition of emotion, as emotion is not seen purely as a physiological response, but instead as a social tool (Darwin, 1872; Frijda, 1986). The Behavioural Ecological View of facial expressions (BECV; Fridlund, 1994, 2017) argues that facial expressions indicate the sender's most likely future behaviours and thus function as important social signals in social interaction. As such, both primary and secondary emotions *can* be associated with specific, readable and recognisable, facial signals and facial behaviours, as it is not the emotion per se that is being transmitted, but instead the potential social action (Fridlund, 2017; Waller et al., 2016). Finally, in a recent set of studies, Crivelli et al (2017) compared different categorisation method, using photographs of spontaneous facial behaviours carrying specific emotional displays. Comparing free label, dimensional ratings, and forced-choice categorisation, they showed that Trobrianders categorised facial displays more reliably and accurately when using dimensional ratings, i.e., defining the emotional states based on valence and arousal (Crivelli et al, 2017). During social interactions, it is indeed easier to understand whether an individual is experiencing a positive or negative (valence) or a mild or intense emotion (arousal; Fridlund 1994; Crivelli et al, 2017) than to identify the exact emotional state the individual is experiencing by looking at one's facial displays.

Nevertheless, even though the content of the information carried out by faces remains debated, the way faces with 'emotional' content are processed has been shown to differ from how neutral faces are processed. When looking at a face in order to extract socially relevant information (i.e., gender, ethnicity, age, familiarity), research showed that individuals from different cultures exhibited different fixation patterns, but those patterns were consistent across tasks (e.g., learning, recognition, categorisation). Western White observers focussed more attention to the eye region and partially to the mouth, whereas East Asian observers focussed more on the central region of the face (Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Kelly et al., 2011; Rodger, Kelly, Blais, & Caldara, 2010). However, decoding studies revealed

different strategies when looking at emotional faces of primary emotions. Researchers looked at either eye movement and fixation patterns across the face (looking at where people looked on a face for sources of information) or at the information extracted and processed leading to accurate recognition (i.e., looking whether the mouth or the eyes hold more relevant/accurate information; Beaudry, Roy-Charland, Perron, Cormier, & Tapp, 2014; Eisenbarth & Alpers, 2011; Vaidya, Jin, & Fellows, 2014). When fixation patterns were investigated, the results usually indicated a higher proportion of fixation on the eye areas and the mouth/lips (Beaudry et al., 2014; Blais, Fiset, Roy, Saumure Régimbald, & Gosselin, 2017; Eisenbarth & Alpers, 2011; Jack et al., 2009; Vaidya et al., 2014). On the one hand, when looking at the face regions fixated during a classic labelling task of basic emotions, the relative proportion of fixation on the different regions was not influenced by the emotion being processed (Jack et al., 2009; Vaidya et al., 2014). The patterns tend to become more clearly differentiated across emotional expressions when participants are presented with more subtle expressions (Vaidya et al., 2014), when the task requires looking for specific emotions (Schurgin et al., 2014), or when presented with dynamic stimuli (Blais et al., 2017). On the other hand, when looking at the information carried by each face region, different patterns have been identified for the six basic emotions: the mouth holds more useful information for the identification of happiness and surprise; the eyes are used to recognise anger and fear; the nose region is the most useful for disgust; and finally the eyebrows and mouth are utilised to identify sadness (Smith, Cottrell, Gosselin, & Schyns, 2005). Those areas are not necessarily fixated for longer but they contain essential information regarding the emotion expressed on the face; when hidden, participants struggled to identify the associated emotion (e.g., lower accuracy in labelling happy faces when the mouth was covered).

Cultural differences in fixation patterns have been previously identified, where people from East-Asian cultures focussed more on the information carried out by the eyes when labelling emotional expressions (Jack et al., 2009), and spent more time looking at the eyes than

any other areas of the face (Jack et al., 2009). In contrast, Western observers distributed their attention evenly across the face, looking at each face region for a similar amount of time (Jack et al., 2009). Moreover, cultural differences have been identified in categorisation tasks, with East-Asian individuals biasing their response to ambiguous emotional stimuli towards less negative or socially-threatening emotions (Matsumoto, 1992; Russell, 1991; Schyns, 1998; Wierzbicka, 1992), mislabelling “fear” and “disgust” for respectively “surprise” and “anger” (Jack et al., 2009) when compared with Western individuals. This could be due to East-Asian individuals sampling information mainly from the eyes, leading to poor dissociation between fear/disgust and surprise/anger emotional expressions. For example, when looking at the prototypical displays identified by Ekman and colleagues (Ekman & Friesen, 1969), it is difficult to differentiate fear/surprise and anger/disgust using the eyes only. The eyes in both fear and surprise present AUs 1+2 (AU1: Inner Brow Raiser; AU2: Outer Brow Raiser; FACS) and the eyes in anger and disgust both exhibit a searching look due to either AU 4 (Brow Lowerer) or AU 9 (Nose Wrinkle; FACS). Cultural differences in eye movements might thus result from cultural specificity in expressing and perceiving emotional signals, where cultural emotional subduction leads to focussing more on the eyes than the mouth (Jack et al., 2009; Yuki, Maddux, & Masuda, 2007).

Therefore, one way to examine whether facial movements contain information that can be considered ‘emotional’ is to examine patterns of eye movements and investigate the informative value of the different parts of the face. However, few studies have looked at the fixation patterns associated with the recognition of secondary emotions (Roy, Blais, Fiset, Rainville, & Gosselin, 2015; Schurgin et al., 2014), and to the best of our knowledge, only one study has looked at fixation patterns associated with guilt (Yu et al., 2017). Guilt is an emotional and cognitive experience arising when someone feels that they did something wrong. It could have evolved due to its potential adaptive function, within social interaction, of encouraging social behaviours towards and from others. Guilt is classified as a moral, self-conscious emotion,

along with pride, shame and embarrassment (Haidt, 2003), and is one of the most social, other-oriented emotions that people experience throughout life (J. Carroll, 1985; Cryder et al., 2012). People often state that they can detect a feeling of guilt in others (Weisman, 2014), but until recently, and with some notable tries (Keltner & Buswell, 1996), no behavioural signals had been reliably associated with the experience of guilt. Our previous study examined what people saw when identifying guilt on the face of another (Julle-Danière et al., under review; Chapter 2). We induced guilt experimentally, eliciting patterns of movement that were associated with both the participants' subjective feelings of guilt and judges' impressions of their guilt (Julle-Danière et al., under review; Chapter 2). In their study, Yu et al. (2017) investigated where a guilty person would look on the face of their victim, revealing more fixation on the nose region that can be seen as an indication of social avoidance. The guilty person will still look towards their victims but will avoid eye contact, which is a behavioural difference between guilt and shame: a shameful person tend to avoid any form of contact by looking down (Keltner, 1995; Keltner & Buswell, 1996). However, where people look and whether some face regions hold more meaning than others to accurately recognise guilt remains unknown.

### **Present Investigation**

In this study, we investigated the scanning patterns associated with the recognition of guilty expressions. We combined a classic forced-choice categorisation task with the measure of ocular movements in people from two different cultural groups. We looked at scanning patterns used to identify guilt compared to those used when viewing basic emotions and tested for cultural differences in eye movements. We built on previous research on emotion recognition using more ecologically valid stimuli: participants were presented with spontaneous dynamic facial expressions of a secondary emotion. The first research question is whether facial expressions displaying a secondary emotion are processed as if they convey emotional information, similar to primary emotions, or as a neutral face. The second research question is, if facial expressions of guilt are indeed processed as presenting emotional information, whether guilt displays are



interpreted as a distinct category of emotional signal (Barrett et al., 2007; Ekman & Cordaro, 2011) or more as a social tool (Barrett et al., 2019; Fridlund, 2017; Waller et al., 2016). If guilt is identified in terms of distinct emotion, we should identify a fixation pattern unique for guilty displays, differing from the patterns previously identified for each primary emotion. However, if, as suggested in Chapter 1, secondary emotions are a mix of several primary emotions, the fixation pattern identified for guilty faces should overlap with some of the previously identified patterns. This would also support the view that faces are processed as a social tool, and the valence and arousal of the emotional expression help observers differentiating between potential social outcomes (see Chapter 4). Finally, the third research question is whether there are cultural differences in the processing of guilty displays. Our previous study (Julle-Danière et al., under review; Chapter 2) showed that self-reported guilt was associated with upper lip movement and neck touching, also present in anger and disgust, whereas perceived guilt was most closely associated with frowning and face touching, also present in anger (and disgust to an extent - FACS; Ekman & Friesen, 1978; Ekman & Keltner, 1970). Fixating the eyes only, as East Asian observers have been found to do, could lead to cultural differences in labelling facial displays of guilt.

## 5.2. Methods

### Judges

Thirty-four Western (WC; 31 Europeans, 25 female, 9 male) judges and 26 Asian (EA; 14 female, 11 male; see Table 5.1 for details) judges took part in the study. Participants' cultural groups were self-defined as they could report the ethnic group they identified with (see *Demographics* section below for details). Participants' age ranged between 18 and 74 years, with a median of 25.5 and an average of 31.8 years ( $SD = 14.04$  years; WC:  $M = 33.32$ ,  $SD = 15.89$ ; EA:  $M = 28.5$ ,  $SD = 10.26$ ). Twenty-three Asian observers were Asian nationals, one had a double

nationality, two were British nationals, and one was European national. Asian observers had an average UK residence of 3.33 years ( $SD = 4.89$ ) but self-identified more as belonging to an Asian cultural background; Western observers had an average UK residence of 29.11 years ( $SD = 17.70$ ). All judges had normal or corrected vision, gave written informed consent, and received £10 for participating. The projects were reviewed and approved by the Science Faculty Ethics Committee (SFEC) from the University of Portsmouth (Appendix 4).

**Table 5.1. Participants demographic information**

<b>Nationality</b>	<b>Count</b>	<b>Average time spent in UK</b>
<b>Asian</b>	<b>26</b>	<b>3.33</b>
Bahraini	1	1.00
Bangladeshi	1	1.66
British	2	14.00
British Pakistani	1	16.08
Chinese	5	1.10
Indonesian	9	1.10
Japanese	1	0.58
Malaysian	4	3.69
Myanmar	1	1.00
Thai	1	8.00
<b>Western</b>	<b>34</b>	<b>29.11</b>
British	29	33.41
Dominican	2	0.12
Dutch	1	14.25
Italian, Srilankam	1	4.00
Spanish	1	2.33
<i>Grand Total</i>	60	17.94

*N.B. Time spent in the UK presented in years*

### **Demographics and acculturation measures**

Participants were asked to indicate the ethnic group they identify themselves as belonging to. They were also asked to indicate their birth country, as well as the birth countries of their parents and grandparents. The participants that reported being born in an Asian country were asked to fill out a four-item scale measuring the degree of *social contact with members of the*

*host culture* (see Supplementary Material Chapter 4, Asian questionnaire; De Leersnyder et al., 2011). Participants reporting being born elsewhere were asked to fill out a four-item questionnaire measuring the degree of exposure to and engagement with non-Western cultures (see Supplementary Material Chapter 4, Western questionnaire; Delis et al., 2016). We excluded three Westerners that spent over six months in Asian countries and one British national born in Indonesia, leaving us with 30 Western participants.

## **Stimuli**

Stimuli consisted of 36 videos displaying eight FACS-coded facial expressions (posed basic emotions: “angry”, “disgust”, “fear”, “happy”, “sad”, “surprise”, induced secondary emotion: “guilt”, and spontaneous expression not judged as guilt: “control”). Ethnicities were equally distributed for each expression (with the exception of fear). All stimuli videos were selected from a previous study where guilt had been induced experimentally and basic emotions posed (Julle-Daniere et al, under review; Chapter 2).

The original 280 videos (211 basic emotions, 69 guilt) were used in a screening study where we collected judgement and impressions, allowing us to select 12 emotion videos (2 cultures x 6 basic emotions) and 12 guilt videos of people reporting the highest level of guilt (felt guilt) and that received the highest rating of guilt (judged guilt; Julle-Daniere et al, under review; Chapter 2). We combined those ratings to facial movements produced and associated with guilty feelings to narrow down our selection to 12 good examples (see Chapter 4 for details). We also selected 12 videos displaying people that reported a medium to low guilt (1 to 3 out of 5 on PANAS self-reported guilt) and that were judged as experiencing an average to low level of guilt (under 40% of judged guilt; see Chapter 4 for details) to have control videos presenting spontaneous facial expressions differing from the displays of guilt. A t-test revealed that those “control” participants reported a similar level of felt guilt than did “guilt” participants ( $t(19.56) = -0.55; p$

= 0.591), but were judged as presenting a significantly lower level of guilt than did the “guilt” participants ( $t(21.39) = -3.46$ ;  $p = 0.0023$ ). We cropped all videos to eight seconds around the apex of the expressions displayed. The six basic emotions and the control displays were included in this study in order for the results presented in this paper to represent the information needed to dissociate guilt from other displays.

## **Apparatus**

The experiment took place in a dimly lit room, using a Samsung laptop (17” LCD monitor, resolution 1024x768, running under Windows 10) connected to an eye-tracking device (Tobii X2-30 Compact; remote system, i.e. not head-mounted) and randomly presenting the 36 videos using the Tobii Studio software. Both monitor and eye tracker were about 60cm from the observers (facial stimuli visual angle:  $10^\circ \times 14^\circ$ ). Eye movements were recorded at a sampling frequency of 30 Hz; data were collected from both eyes (binocular), with an average gaze accuracy of  $<0.5^\circ$  and a gaze precision of  $<0.45^\circ$  (measured under ideal conditions). The eye tracker’s accuracy and precision are conserved for blinks and moderate head movements.

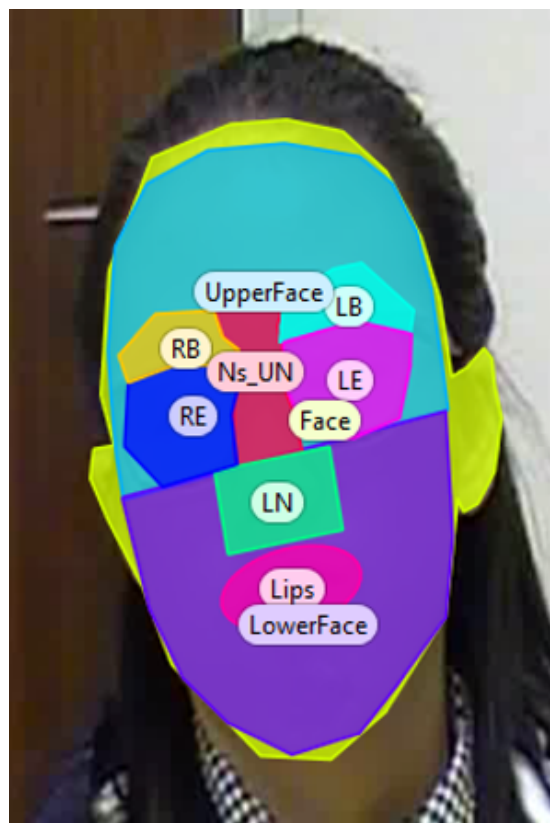
## **Procedure**

After giving written consent and filling out the demographic questionnaires, participants were seated in front of the eye-tracking equipment. Before starting the tasks, we performed a 5-point calibration of the eye tracker for each participant using Tobii Studio. Once the calibration was satisfactory, participants started the categorisation task. Every participant watched the 36 videos. After each video, they were asked to indicate on a tablet provided which main emotion they thought the person in the video was feeling, and whether they thought the person in the video was experiencing other emotions alongside the main one identified previously. They did so by performing two eight-alternative forced-choice tasks, selecting one of the following labels:

“anger”, “fear”, “disgust”, “happiness”, “surprise”, “sadness”, “guilt”, or “none of the above”. No contextual information was provided with the videos. There was no time pressure to complete the categorisation tasks; participants could move on to the next video when they wanted to, allowing them to take breaks between stimuli if they so wished to.

### Face Regions

For each video stimulus, ten face regions were defined, following the template presented in Figure 5.1. Those areas of interest were identified from previous eye-tracking studies looking at the categorisation of emotional faces (Blais et al., 2017; Jack et al., 2009; Schurgin et al., 2014), identifying the eyes (eyebrows and eyes), the nose (upper and lower parts) and the lips as the main zones of interest.



**Figure 5.1. Illustration of 10 facial regions of interest (ROIs).**

RB: right brow; RE: right eye; LB: left brow; LE: left eye; Ns\_UN: nasion and upper nose; LN: lower nose; UpperFace: from the forehead to the middle of the nose; LowerFace: from the middle of the nose to the chin; Face: the whole face.

## Statistical Analysis

### Task performance

We used a generalised linear mixed model approach (GLMM) to examine the decoding of facial expressions of emotions by different cultures. We looked at the impact of culture on mean categorisation accuracy, investigating whether the culture of the judges or the culture of the stimulus influenced the accuracy of the judges. Mean categorisation accuracy of the judge by emotion categories was set as our response variable, and the cultural congruency (WC-EA, WC-WC, EA-EA, and EA-WC) and the facial expression (8 x facial expressions) were set as predictors. As our dataset considered ratings of videos by multiple judges, we include the identity of the judge, and the stimulus individual, as random factors in the model to control for pseudoreplication (Hurlbert, 1984; Waller et al., 2013). We fitted the GLMM using the function `lmer` provided by the packages `lme4` and `lmerTest` for RStudio Version 0.99 for R 287 version 3.1.3 (Bates et al., 2014). The GLMM was fitted with a Gaussian error structure and estimated using REML. To investigate how judges labelled each facial display, we looked at the choices made by judges when looking at European (WC) or Asian (EA) faces for the different emotion types. In order to know which label was preferentially selected by judges for each facial display, we computed a chi-square and a cross table to investigate the distribution of choices over the 8 facial displays presented.

### Eye-tracking data

To investigate the information used to categorise facial expressions, i.e., where do judges look most when looking at each stimulus set, we first ran correlations between the proportion of movements produced (i.e., the number of AUs displayed in a given region of the face) in the upper face and the number of fixations made to the upper face ( $\frac{\text{number of fixation made to the upper face}}{\text{total number of fixation to the whole face}}$ ); same for the lower face). We wanted to see

whether participants were looking at specific areas for the information they presented or simply because their eyes were attracted to moving areas.

Then, we examined the frequency of fixations (number of times they looked at specific areas) to the different face regions for both judge groups (European and Asian): the frequency of fixation was our response variable, and the culture of the judge (2 x culture of judges), the culture of the stimulus' face (2 x culture of face), the cultural congruency (EU-EA, EU-EU, EA-EA, and EA-EU), and the facial expression (8 x facial expressions) were set as predictors. We ran two sets of analyses in order to get a better understanding of the fixation pattern displayed by judges for the different emotional expressions. First, we looked at two face regions only: "upper" and "lower" face; second, we looked at four different face regions: "Eyes" (left and right eyebrows and eyes), "Upper Nose", "Lower Nose", and "Lips" (based on Jack et al., 2009).

### 5.3. Results

#### **Task performance**

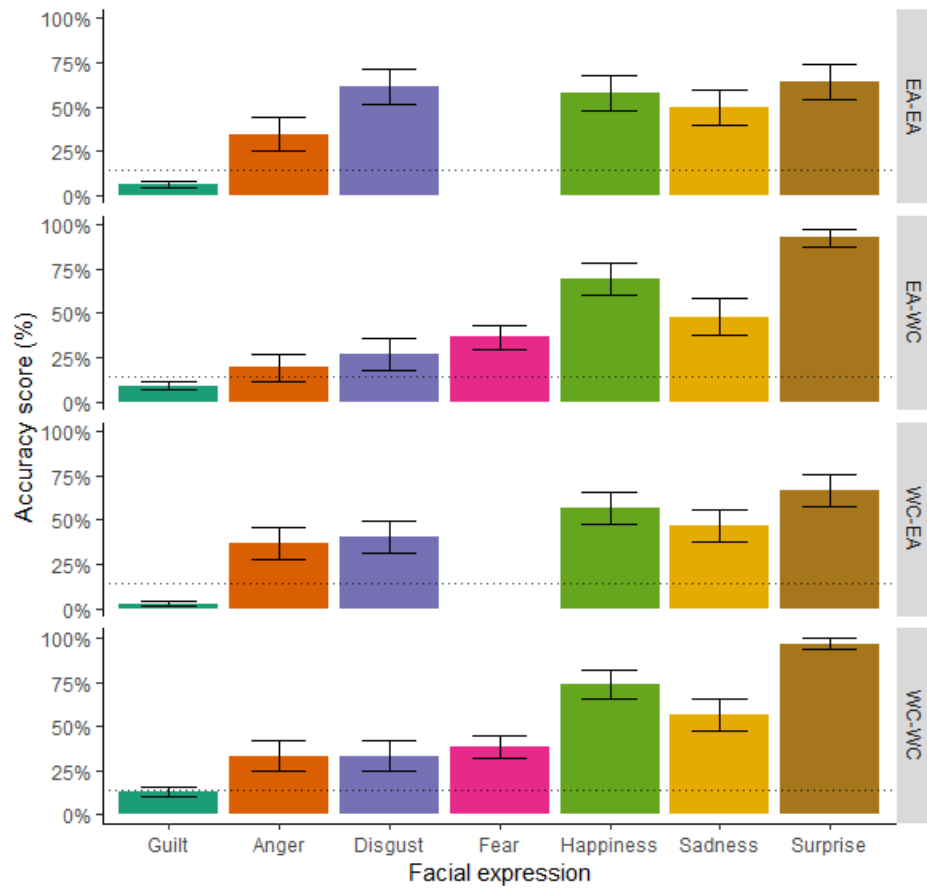
We found a significant main effect of facial expression: guilt was less accurately labelled than all other emotional displays ( $p < 0.001$ ), including control displays ( $\beta = 2.82e-1$ ;  $SE = 0.6.68e-2$ ;  $p = 0.00016$ ;  $r^2 = 0.28$ ). All participants, regardless of their culture, accurately labelled the control displays as "None of the above"; we thus decided to re-run the GLMM without the control displays in order to identify the potential cultural difference between guilt displays and displays of primary emotions.

We looked at the impact of cultural congruency (WC-EA, WC-WC, EA-EA, and EA-WC) and facial displays (8 x facial expressions) on categorisation accuracy. We found a main effect of

facial expression: the guilt display was less accurately labelled than the other displays ( $p < 0.001$ ). Additionally, we found significant facial displays x congruency interactions. Post-hoc Bonferroni comparisons revealed that overall all judges categorised guilt with the same accuracy regardless of their culture and of the culture of the video; however, neither WC judges nor EA judges looking at WC face differentiated anger, disgust, or guilt displays from each other ( $ps = 1$ ). When looking at EA faces, EA judges did not differentiate anger from guilt displays ( $p = 0.096$ ). When looking at EA faces, WC judges accurately differentiated guilty displays from all other displays ( $ps < 0.0016$ ; Figure 5.2). From this, it appeared that WC faces displayed anger, disgust and guilt in such a way that judges did not differentiate between the 3 displays. Moreover, EA judges were not able either to differentiate anger from guilt when looking at EA faces, whereas WC judges did differentiate EA guilt displays from all other EA displays. EA judges thus had the lowest differentiation accuracy when looking at either same-culture or other-culture faces.

Further analysis revealed that judges made mixed attributions when looking at guilty faces: judges mainly selected "Fear" ( $p < 0.01$ ) for guilt displays more often than the other labels. Inversely, "Happiness" was rarely selected for guilt displays ( $p < 0.001$ ; see Table 5.2 for details).





**Figure 5.2. Accuracy scores presented for each facial expression in each cultural group**  
 Top 2 graphs present the results for EA judges looking at EA (top) or WC (bottom) faces;  
 bottom 2 graphs represent the results for WC judges looking at EA (top) or WC (bottom)  
 faces (*captions read: judge-face*).

Table 5.2. Cross Table results showing choices made for each facial expression

Facial expression	Label selected									
	Anger	Disgust	Fear	Guilt	Happiness	None of the above	Sadness	Surprised		
Anger	Count	35	18	10	20	1	12	9	7	
	Expected values	6.611	8.222	14	8.833	9.389	20.278	13.056	31.611	
	Row percent	<b>31.25%</b>	16.07%	8.93%	17.86%	0.89%	10.71%	8.04%	6.25%	
	Std residual	<b>11.041</b>	<b>3.41</b>	-1.069	<b>3.757</b>	-2.738	-1.838	-1.122	-4.377	
Disgust	Count	2	45	8	4	2	12	15	24	
	Expected values	6.611	8.222	14	8.833	9.389	20.278	13.056	31.611	
	Row percent	1.79%	<b>40.18%</b>	7.14%	3.57%	1.79%	10.71%	13.39%	21.43%	
	Std residual	-1.793	<b>12.826</b>	-1.604	-1.626	-2.411	-1.838	0.538	-1.354	
Fear	Count	3	9	42	7	2	13	13	23	
	Expected values	6.611	8.222	14	8.833	9.389	20.278	13.056	31.611	
	Row percent	2.68%	8.04%	<b>37.50%</b>	6.25%	1.79%	11.61%	11.61%	20.54%	
	Std residual	-1.404	0.271	<b>7.483</b>	-0.617	-2.411	-1.616	-0.015	-1.532	
Guilt	Count	52	42	111	52	38	116	79	181	
	Expected values	39.608	49.26	83.875	52.921	56.25	121.486	78.217	189.384	
	Row percent	7.75%	6.26%	<b>16.54%</b>	7.75%	5.66%	17.29%	11.77%	26.98%	
	Std residual	<b>1.969</b>	-1.034	<b>2.962</b>	-0.127	-2.433	-0.498	0.089	-0.609	
Happiness	Count	5	2	5	0	72	8	10	10	
	Expected values	6.611	8.222	14	8.833	9.389	20.278	13.056	31.611	
	Row percent	4.46%	1.79%	4.46%	0.00%	<b>64.29%</b>	7.14%	8.93%	8.93%	
	Std residual	-0.627	-2.17	-2.405	-2.972	<b>20.434</b>	-2.727	-0.846	-3.844	
None of the above	Count	16	24	70	56	49	183	52	224	
	Expected values	39.785	49.48	84.25	53.158	56.501	122.029	78.566	190.231	
	Row percent	2.37%	3.56%	10.39%	8.31%	7.27%	<b>27.15%</b>	7.72%	33.23%	
	Std residual	-3.771	-3.622	-1.552	0.39	-0.998	<b>5.519</b>	-2.997	2.448	
Sad	Count	1	5	6	20	1	12	56	10	
	Expected values	6.552	8.149	13.875	8.754	9.305	20.097	12.939	31.329	
	Row percent	0.90%	4.51%	5.41%	18.02%	0.90%	10.81%	<b>50.45%</b>	9.01%	
	Std residual	-2.169	-1.103	-2.114	<b>3.801</b>	-2.723	-1.806	<b>11.971</b>	-3.811	
Surprise	Count	5	3	0	0	4	9	1	90	
	Expected values	6.611	8.222	14	8.833	9.389	20.278	13.056	31.611	
	Row percent	4.46%	2.68%	0.00%	0.00%	3.57%	8.04%	0.89%	<b>80.36%</b>	
	Std residual	-0.627	-1.821	-3.742	-2.972	-1.759	-2.504	-3.336	<b>10.385</b>	

Row percent in red and bold represent the main choice made per facial expression; Std residual in red show significant values.

Total number of judgements made facial expressions associated with basic emotion: 2 videos \* 56 participants = 112

Total number of judgements made facial expressions associated with guilt/control: 12 videos \* 56 participants = 672

## Eye-tracking data

### Upper vs Lower Face

First, the number of movements produced in the upper or lower face did not correlate with the number of fixations made respectively to those regions (upper face:  $r = -0.26$ ,  $p = 0.230$ ; lower face:  $r = 0.30$ ,  $p = 0.166$ ). Thus, any fixation pattern that we found does not result from a response to movement: judges looked at specific areas of the face because of preferences to sample those specific areas and not because the judges' eyes were attracted to movement.

Second, we found a significant main effect of the face region, regardless of the expression type stimuli: judges spent more time looking at the upper face than they did at the lower face ( $\beta = 1.29e-01$ ;  $SE = 2.34 e-02$ ;  $p < 0.001$ ). Moreover, Bonferroni posthoc correction revealed that judges spent less time looking at the lower face when looking at displays of sadness than when looking at displays of guilt ( $p = 0.0198$ ). Judges looked at the upper face a similar amount of time in guilt displays compared to other emotional displays (all  $p > 0.111$ ).

### Four face regions

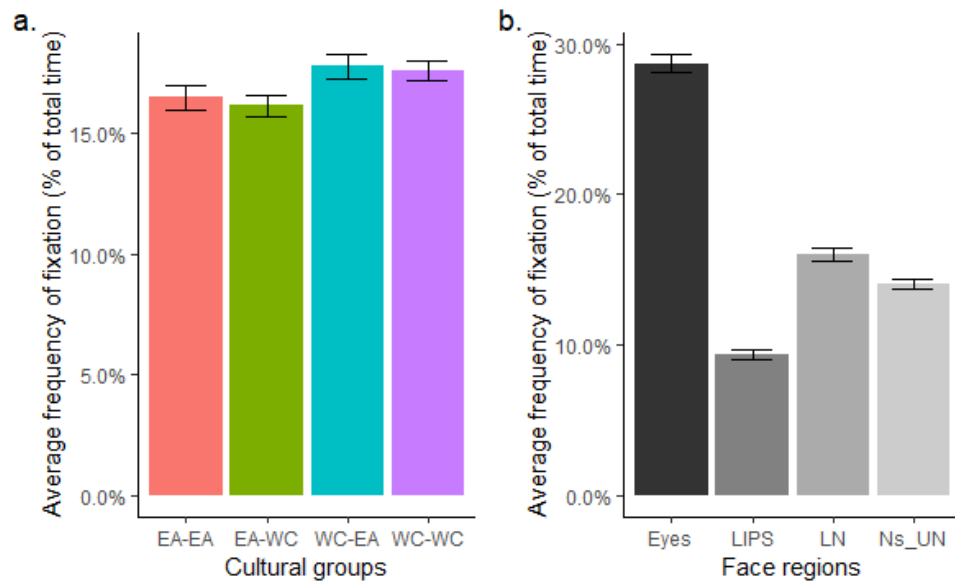
**Main effects.** In this second model, we found a significant main effect of cultural congruency where WC judges looked at EA faces for longer than EA judges (see Table 5.3; see Figure 5.3.a). Multiple comparisons also revealed that WC judges looked at EA face for longer than EA judges looking at WC faces (see Table 5.3). We found a main effect of facial expression: judges looked for longer at control and disgust displays than they did at guilt displays (see Table 5.3). We also found a main effect of face regions: judges looked more at the eye area than any other regions (see Table 5.3; see Figure 5.3.b). Finally, we found some

significant interactions between congruency and some face regions, as well as between the facial expression and the face regions (see below for details).

**Table 5.3. Linear mixed model for 4 face regions looking at cultural congruency between judges and videos**

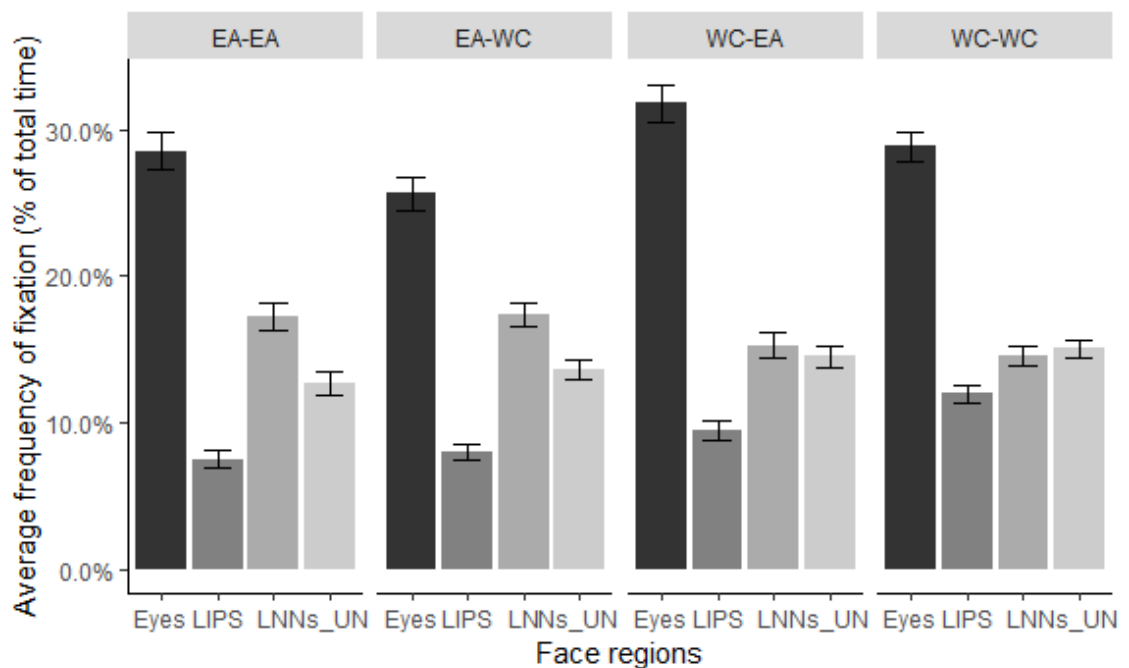
Fixed effects	Estimate	Std. Error	df	t value	Pr(> t )
(Intercept)	2.56E-01	1.79E-02	9.63E+01	14.297	< 2e-16 ***
<b>Main effects</b>					
CongruencyEA-EU	-3.33E-02	1.72E-02	8.43E+01	-1.935	0.056342 .
<b>CongruencyEU-EA</b>	<b>3.26E-02</b>	<b>1.65E-02</b>	<b>1.82E+02</b>	<b>1.974</b>	<b>0.049871 *</b>
CongruencyEU-EU	-1.09E-03	2.05E-02	1.21E+02	-0.053	0.957588
EmotionAnger	2.78E-02	3.30E-02	4.65E+01	0.843	0.403766
<b>EmotionControl</b>	<b>5.28E-02</b>	<b>1.77E-02</b>	<b>4.65E+01</b>	<b>2.989</b>	<b>0.004464 **</b>
<b>EmotionDisgust</b>	<b>1.03E-01</b>	<b>3.30E-02</b>	<b>4.65E+01</b>	<b>3.122</b>	<b>0.003089 **</b>
EmotionFear	6.74E-02	3.39E-02	4.65E+01	1.989	0.052545 .
EmotionHappiness	4.82E-02	3.30E-02	4.65E+01	1.46	0.151089
EmotionSad	-1.06E-02	3.30E-02	4.65E+01	-0.321	0.749338
EmotionSurprise	7.98E-03	3.30E-02	4.65E+01	0.242	0.810208
<b>FaceRegionLIPS</b>	<b>-1.93E-01</b>	<b>1.47E-02</b>	<b>7.94E+03</b>	<b>-13.151</b>	<b>&lt; 2e-16 ***</b>
<b>FaceRegionLN</b>	<b>-8.51E-02</b>	<b>1.47E-02</b>	<b>7.94E+03</b>	<b>-5.81</b>	<b>6.51E-09 ***</b>
<b>FaceRegionNs_UN</b>	<b>-1.37E-01</b>	<b>1.47E-02</b>	<b>7.94E+03</b>	<b>-9.349</b>	<b>&lt; 2e-16 ***</b>
<b>Congruency*FaceRegion</b>					
<b>CongruencyEA-EU:FaceRegionLIPS</b>	<b>4.14E-02</b>	<b>1.71E-02</b>	<b>7.94E+03</b>	<b>2.417</b>	<b>0.015686 *</b>
CongruencyEU-EA:FaceRegionLIPS	-1.30E-02	1.68E-02	7.94E+03	-0.773	0.43932
<b>CongruencyEU-EU:FaceRegionLIPS</b>	<b>4.88E-02</b>	<b>1.66E-02</b>	<b>7.94E+03</b>	<b>2.944</b>	<b>0.003253 **</b>
<b>CongruencyEA-EU:FaceRegionLN</b>	<b>3.43E-02</b>	<b>1.71E-02</b>	<b>7.94E+03</b>	<b>2.002</b>	<b>0.045364 *</b>
<b>CongruencyEU-EA:FaceRegionLN</b>	<b>-5.21E-02</b>	<b>1.68E-02</b>	<b>7.94E+03</b>	<b>-3.109</b>	<b>0.001885 **</b>
CongruencyEU-EU:FaceRegionLN	-2.61E-02	1.66E-02	7.94E+03	-1.575	0.115373
<b>CongruencyEA-EU:FaceRegionNs_UN</b>	<b>4.39E-02</b>	<b>1.71E-02</b>	<b>7.94E+03</b>	<b>2.565</b>	<b>0.010341 *</b>
CongruencyEU-EA:FaceRegionNs_UN	-1.44E-02	1.68E-02	7.94E+03	-0.859	0.39043
CongruencyEU-EU:FaceRegionNs_UN	2.61E-02	1.66E-02	7.94E+03	1.576	0.115029
<b>Emotion*FaceRegion</b>					
EmotionAnger:FaceRegionLIPS	-4.74E-02	2.63E-02	7.94E+03	-1.799	0.072028 .
EmotionControl:FaceRegionLIPS	-2.16E-02	1.41E-02	7.94E+03	-1.533	0.125194
<b>EmotionDisgust:FaceRegionLIPS</b>	<b>-5.97E-02</b>	<b>2.63E-02</b>	<b>7.94E+03</b>	<b>-2.265</b>	<b>0.023557 *</b>
<b>EmotionFear:FaceRegionLIPS</b>	<b>-8.73E-02</b>	<b>2.70E-02</b>	<b>7.94E+03</b>	<b>-3.235</b>	<b>0.001223 **</b>
EmotionHappiness:FaceRegionLIPS	-2.97E-02	2.63E-02	7.94E+03	-1.128	0.259299
EmotionSad:FaceRegionLIPS	-2.71E-02	2.63E-02	7.94E+03	-1.03	0.30306
EmotionSurprise:FaceRegionLIPS	-5.58E-03	2.63E-02	7.94E+03	-0.212	0.832374
EmotionAnger:FaceRegionLN	-5.11E-02	2.63E-02	7.94E+03	-1.94	0.052443 .
<b>EmotionControl:FaceRegionLN</b>	<b>-6.16E-02</b>	<b>1.41E-02</b>	<b>7.94E+03</b>	<b>-4.378</b>	<b>1.21E-05 ***</b>
<b>EmotionDisgust:FaceRegionLN</b>	<b>-5.50E-02</b>	<b>2.63E-02</b>	<b>7.94E+03</b>	<b>-2.089</b>	<b>0.036738 *</b>
<b>EmotionFear:FaceRegionLN</b>	<b>-6.21E-02</b>	<b>2.70E-02</b>	<b>7.94E+03</b>	<b>-2.302</b>	<b>0.021379 *</b>
EmotionHappiness:FaceRegionLN	2.57E-02	2.63E-02	7.94E+03	0.977	0.328482
EmotionSad:FaceRegionLN	-9.88E-03	2.63E-02	7.94E+03	-0.375	0.707539
EmotionSurprise:FaceRegionLN	-1.23E-02	2.63E-02	7.94E+03	-0.468	0.639959
EmotionAnger:FaceRegionNs_UN	-1.87E-02	2.63E-02	7.94E+03	-0.712	0.476698
<b>EmotionControl:FaceRegionNs_UN</b>	<b>-4.79E-02</b>	<b>1.41E-02</b>	<b>7.94E+03</b>	<b>-3.404</b>	<b>0.000668 ***</b>
EmotionDisgust:FaceRegionNs_UN	-4.35E-02	2.63E-02	7.94E+03	-1.651	0.098736 .
<b>EmotionFear:FaceRegionNs_UN</b>	<b>-7.18E-02</b>	<b>2.70E-02</b>	<b>7.94E+03</b>	<b>-2.66</b>	<b>0.007832 **</b>
EmotionHappiness:FaceRegionNs_UN	-4.19E-02	2.63E-02	7.94E+03	-1.59	0.111898
EmotionSad:FaceRegionNs_UN	2.90E-02	2.63E-02	7.94E+03	1.1	0.27147
EmotionSurprise:FaceRegionNs_UN	-2.09E-03	2.63E-02	7.94E+03	-0.079	0.936902

Formula: Count ~ (Congruency + Emotion) \* FaceRegion + (1 | Participant) + (1 | Video)



**Figure 5.3. a. Average frequency of fixation for each congruent group; b. Average frequency of fixation to each face region.**

**Interaction cultural congruency x face region.** Post-hoc comparison on the cultural congruency and face regions showed that the eyes, the lower nose, and the upper nose were fixated evenly by all judges regardless the culture of the face ( $p > 0.322$ ). However, WC judges fixated WC lips more than EA judges fixated EA lips ( $p = 0.0112$ ; see Figure 5.4).



**Figure 5.4. Average frequency of fixation to each face region depending on the cultural group**

**Interaction facial expression x face region.** Finally, posthoc comparison looking at the interaction between facial expressions and face regions showed that regardless of the culture of video or judge, judges looked at the eye region the most for all facial displays ( $p < 0.001$ ; see Figure 5.5), except happiness where eyes and lower nose were fixated for similar amount of time ( $p = 0.968$ ). Moreover, the eyes were fixated less in guilt displays than in disgust ( $p < 0.001$ ) and control displays ( $p < 0.001$ ). The time spent looking at the lips was not different for guilt compared to other emotional displays (all  $ps > 0.615$ ) and the lips were the least fixated region. Judges looked at the lower nose less in guilt displays compared to happy displays ( $p = 0.0388$ ). Finally, judges fixated the upper nose region for a similar amount of time for all emotional displays (all  $ps > 0.5$ ). Control displays and guilt displays were thus fixated differently: judges looked at the eyes and mouth for longer in control displays than in guilt displays. Within guilt displays, judges looked the most at the eye region ( $p < 0.001$ ) and the least at the lips ( $p < 0.001$ ); they fixated the lower nose and the upper nose for a similar amount of time ( $p = 1$ ).

Overall, judges spent less looking at the lower face when looking at guilt than sadness; they fixated the eyes less for guilt than for disgust or control displays; they fixated the lower nose less for guilt than for happy displays. Thus, when looking at guilt displays, judges used a fixation pattern that seems very similar to the one used when looking at fear, surprised and angry displays.

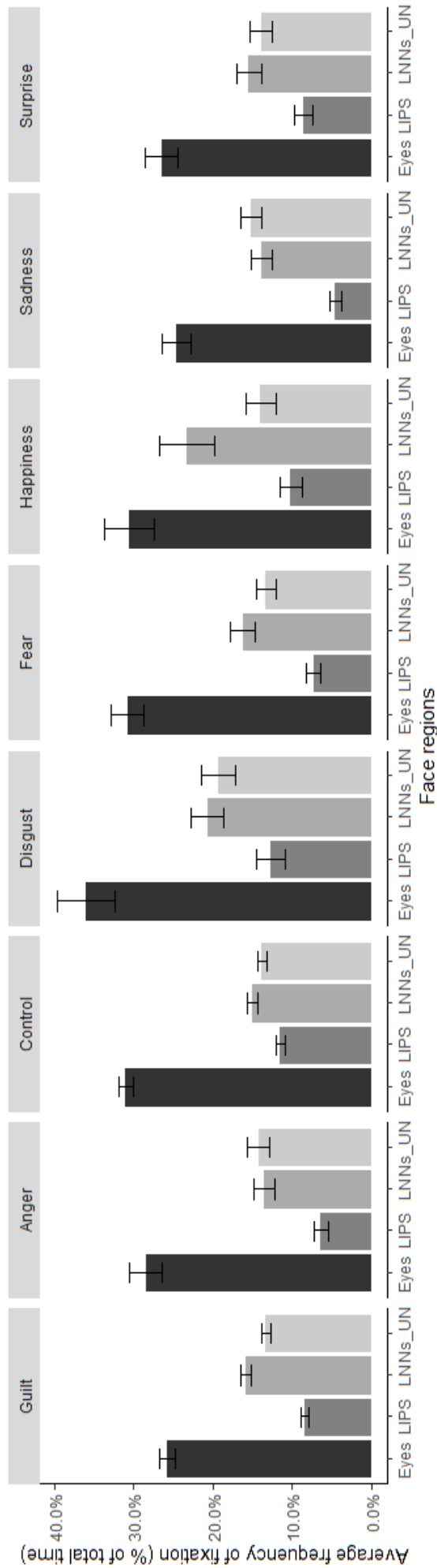


Figure 5.5. Average frequency of fixation to each face region depending on the facial expression presented.

#### 5.4. Discussion

Perceiving facial expressions and being able to understand their meaning in social interaction is a key element of human interactions. In this study, we tried to identify the eye movement patterns used to discriminate the facial expression of guilt, a secondary emotion, from the facial displays of 'basic' emotions. We used dynamic stimuli presenting spontaneous displays associated with genuine feelings of guilt. Our results show that the fixation patterns exhibited by judges looking at guilt displays differed from the patterns previously identified during (non-emotional) face perception. Moreover, the pattern of fixation exhibited when looking at guilty displays significantly differed from the one used when looking at the control displays. This answers our first research question: the displays associated with guilt were perceived as presenting a specific, identifiable, emotional content. Moreover, the fixation pattern used when looking at guilt signals was very similar to the one used when looking at fear, surprised and angry displays. This answers our second research question: the fixation patterns used when looking at guilty displays is not specific to guilt but appears to overlap with some of the previously identified patterns for other emotions. It is thus possible that the emotional display of guilt was perceived as a mix of primary emotions rather than as a unique, specific emotion. Finally, we reported differences between Asian and European observers in the decoding of guilt facial expressions; European observers made more mistakes when looking at Asian guilty expressions and observers from different cultures displayed different scanning strategies. This answers our third and final research question: there are cultural differences in the processing of guilty displays.

To identify a fixation pattern specific to guilty displays, we compared the time spent looking at each face region to establish their relative importance. When sampling guilty faces, observers looked most at the eyes, then fixated the central part of the face (Lower and Upper Nose), and spent the lesser time looking at the lips area. Observers used the same sampling strategy for all



emotional displays presented in this study. However, if the proportion of time spent looking at the different areas was the same for all emotions, the total fixation time to the specific regions varied between emotional displays. This general fixation pattern seems to contradict previous studies showing some variations in fixations for different emotional displays (Beaudry et al., 2014; Schurgin et al., 2014), with the mouth region being most fixated during processing of happiness and the eyes being most involved in the recognition of sadness (Beaudry et al., 2014; Schurgin et al., 2014). However, those differences could be explained by the differences in stimuli types: previous studies tend to have used mainly fixed images depicting emotional states; here, we used either posed-dynamic (basic emotions) or genuine-dynamic (guilt and control) emotional stimuli. It could be that in the case of dynamic faces, observers focus mainly on the eyes of the person they are interacting with, sampling other areas as a secondary measure only. Finally, the fixation on the eye region could suggest that guilt is most easily identified from the eyes as it was suggested in a previous study (Yu et al., 2017). This is in line with our previous study (Julle-Danière et al., under review; Chapter 2) showing that both self-reported guilt and perceived guilt were associated with eye movements. Indeed, individuals feeling guilty produced more *eye movements* than individuals in a control/neutral situation (see Figure 2.2 and Table 2.3); the perception of guilt by others was most closely associated with *frowning* and face touching (Julle-Danière et al., under review; Chapter 2). It could be interesting to go further and use the Bubbles procedure (Smith et al., 2005), randomly presenting information from various face regions, to identify the relative importance of the eye region in accurately perceiving guilt on a face.

The fixation times to the different face regions displayed for guilt stimuli presented great similarity to the scanning strategy displayed when looking to fearful displayed, and some similarity to the patterns displayed for surprised and angry stimuli. Moreover, guilt was mislabelled for fear significantly more than any other emotional label. It could be that the facial movements produced during the experience of guilt were similar to the ones produced in posed

fearful displays. The common fixation strategy did not allow for a clear dissociation between fearful and guilty stimuli. However, it allowed for a clear differentiation between guilt displays and surprised/angry displays, as surprised and angry stimuli were accurately recognised and differentiated from guilty displays. Thus, when looking at the eye region, guilt stimuli must present similar facial movements to fear displays. Indeed, in our previous study (Julle-Danière et al., under review; Chapter 2) we found that individuals in the guilt condition exhibited a lot of AUs 1+2+5 together (AU1: Inner Brow Raiser, AU2: Outer Brow Raiser, AU5: Upper Lid Raiser; FACS; see Table 2.2, which are also present in the prototypical expression of fear, composed of AUs 1+2+5+20+25+26 (AU20: Lip Stretch, AU25: Lips Part, AU26: Jaw Drop; Ekman & Friesen, 1978).

In line with previous research (Jack et al., 2009; Miellet, Vizioli, He, Zhou, & Caldara, 2013), we found that WC and EA observers deployed different scanning strategies when looking at emotional faces. WC judges fixated more the eye region as well as the lips compared to EA judges; EA judges, however, exhibited a central or global information sampling strategy (Blais et al., 2008; Miellet et al., 2013), characterised by a more central fixation to the lower nose. We also found a difference in scanning pattern based on a same- vs other-race effect: WC looked at the lip area on WC faces more than did EA judges looking at EA faces. The general cultural scanning patterns were emphasised when looking at same-race faces compared to other-race faces. Cultural differences in scanning patterns might result from cultural specificity in expressing emotional signals, where cultural emotional subduction leads to focussing more on the eyes than the mouth (Jack et al., 2009; Yuki et al., 2007).

To conclude, our study revealed that when looking at facial displays associated with secondary emotion, namely guilt, observers exhibited scanning patterns similar to the ones used when looking at facial displays of primary emotions. Observers do not seem to discriminate between emotional displays by using different scanning patterns, but thanks to the information carried out on the face. Some differences in fixation times were found between emotional displays, but the scanning pattern exhibited for guilt displays did not differ from the one used for fear displays, which can explain the mislabelling found here. From here on, it would be interesting to investigate which face region holds the most relevant information when it comes to identifying guilty displays. This could help further our understanding of the perception of secondary emotions, as well as help explain the mislabelling found in this study.

## **6. The social function of the feeling and expression of guilt <sup>5</sup>**

### **Abstract**

Guilt is a social, other-oriented moral emotion that promotes relationship repair and pro-sociality. For example, people can be more lenient towards wrongdoers who display guilt than towards those who do not. However, it is unclear how friendship modulates the social outcomes of guilt. Here we examined the social consequences of guilt in an experimental setting with pairs of friends differing in relationship quality. We found that guilty people were motivated to repair wrongdoing regardless of friendship and observing guilt in others led to a punishment effect. Moreover, a victim of wrongdoing punished close friends who appeared guilty more so than acquaintances. We suggest that guilt has a stronger function between close friends when the costs of relationship breakdown are greater. Relationship context, therefore, is crucial to the evolutionary relevance of moral emotions.

**Keywords:** Facial expression, guilt, emotion, friendship, social consequences, morality

---

<sup>5</sup> Results from this study have been presented with a paper at the International Society for Research on Emotion (Amsterdam 2019). A version of this chapter has also been submitted to *Proceedings of the Royal Society B* and is currently under review.

## 6.1. Background

Humans are uniquely cooperative (Tomasello, 2008) and form crucial short- and long-term relationships between individuals (for review, see Silk, 2005). Such bonds have proven adaptive across species, increasing the fitness of the individuals involved (Christakis & Fowler, 2010; Seyfarth & Cheney, 2012), as well as of their descendants, contributing to their central place in human societies (Christakis & Fowler, 2010). The need for intense cooperation may have provided a particularly powerful selection pressure on the many emotional and communicative behaviours regulating cooperative processes. Such relationships involve cooperative interactions widely separated in time and space (reciprocal, and even delayed, altruism; Hamilton, 1964; Trivers, 1971), based on the memory of past interactions and the emotional load associated with them (Seyfarth & Cheney, 2012; Silk, 2005). The costs associated with delayed altruism are higher than those associated with direct cooperation (there is a risk that the partner will not return the favour; Mayr, 1961; Tinbergen, 1963), necessitating the simultaneous evolution of a control system. Punishment and spite have the potential to maintain cooperative behaviour and ensure equity within the relationship (Henrich & Boyd, 2001; Jensen, 2010; Jensen et al., 2006) by imposing costs on defecting partners. The costs could be the refusal to cooperate at a later time or damage to the defecting partner's reputation at the scale of the whole social group (Haley & Fessler, 2005), leading to wider repercussions against the untrustworthy. To regain the good graces of their partners, defectors need to acknowledge their wrongs, make amends, and even express remorse regarding their wrongdoing (Rosenstock & O'Connor, 2018). Moral emotions are, therefore, intimately linked to our relationships with others, (Haidt, 2003), facilitating the social interactions and important relationships (Keltner & Gross, 1999; Muris & Meesters, 2014).

Guilt is a social, other-oriented emotion that people experience regularly throughout life (J. Carroll, 1985; Cryder et al., 2012; Tangney, 1999). Evidence suggests that guilt has a potentially

positive function within social interaction of stimulating pro-social behaviours towards and from others, promoting actions towards those who have been wronged (Cryder et al., 2012; De Hooge et al., 2007; Ketelaar & Au, 2003). As the quality of relationship and likelihood of repeated interaction between two individuals is likely to affect the consequences of social transgressions (Rosenstock & O'Connor, 2018), however, guilt may perform different roles depending on friendship status. For example, guilt could have a stronger impact within existing friendships compared to new or less close friendships (Hess, Banse, & Kappas, 1995; Wagner & Smith, 1991), as the costs of relationship breakdown are greater. Regardless of any proximate penalties (e.g. punishment), there could be long-term benefits to both parties if the social bond is reinforced. It could thus be advantageous to communicate a feeling of guilt unambiguously within social interaction (Cryder et al., 2012; Julle-Danière et al., under review; Keltner & Buswell, 1996; Weisman, 2014), given that expressing such feelings could maintain the social bond long-term.

Cooperation is a widespread tendency in humans, but some cultural differences have previously been identified (Gächter, Herrmann, & Thöni, 2010; Henrich, Ensminger, et al., 2010; Henrich, Heine, et al., 2010). In collectivistic settings, people develop strong bonds and cooperate mostly with close relatives, investing time and resources in the relationship. In such cultural settings, it is possible that individuals have a clearer signal of guilt to reduce any potential punishment. In line with this hypothesis, cultural differences in the experience (De Leersnyder et al., 2011) and production/perception of emotion (Jack et al., 2016) have been demonstrated by recent work. Whether people experience a specific emotion seems to be related to their cultural norms (Mesquita & Leu, 2007): a specific emotion could be experienced more so in one culture (i.e., anger in European countries) and people in different cultures can experience emotions differently (i.e., anger associated with a feeling of control in European countries but guilt in East-Asian countries). However, these processes do not seem fixed; recent work has shown that people changing cultural environment can 'adapt' to their new culture by experiencing emotional acculturation, slowly beginning to experience and express emotion

according to their new cultural standards (De Leersnyder, 2017). In regards to guilt, individuals in East Asian cultures have been shown to be more sensitive to feelings of guilt, in both themselves and their close relatives, than individuals in European cultures (Cousins, 1989), and also to exhibit greater motivation to minimise disruption to relationship harmony (Kim et al., 2008).

### **Present Investigation**

We investigated the impact of friendship on experiencing and observing guilt within social interaction. In an experimental laboratory study, pairs of friends and pairs of strangers took part in a cooperative task. We then artificially induced guilt in one participant, informing the other participant of their partner's wrongdoing. We examined the influence of friendship, facial expression and culture on the participants' responses using a dictator game (Forsythe, Horowitz, Savin, & Sefton, 1994). We aimed to replicate previous findings showing the pro-social function of guilt and test whether friendship affects how wrongdoers who are perceived as guilty are treated.

## **6.2. Methods**

### **Participants**

Two hundred and eighteen participants were recruited as pairs; participants were given the opportunity to either sign up with a friend or sign up alone and be randomly paired up with a stranger (only 36 participants signed up alone to be opportunisticly paired up with someone else). Participants were recruited based on an opportunistic sampling method and were either part of the University community (students and staff members) or part of the general public. Three participants were excluded from the study as they did not complete all the steps required

and we were, therefore, unable to use their final data for analysis (final sample = 215; 109 pairs: 107 as Player 1 and 108 as Player 2).

**Table 6.1. Participants demographic information**

<b>Nationality</b>	<b>Count</b>	<b>Average time spent in UK</b>
<b>Asian</b>	<b>87</b>	<b>4.12</b>
Bangladeshi	1	1.10
British	14	17.36
British/Indonesian	1	8.00
British/Irish	1	20.50
Bruneian	1	1.58
Chinese	36	0.66
Dominican	1	0.33
Filipino	1	0.42
German	1	1.42
Indian	1	5.00
Indonesian	6	1.58
Japan	1	0.16
Malaysian	15	1.67
Myanmar	1	0.50
Pakistani	1	1.00
Saudi	1	2.00
Spanish	1	7.25
Thai	3	2.56
<b>European</b>	<b>128</b>	<b>17.50</b>
Afghan/British	1	8.25
American	1	0.33
American	2	19.50
American/Dominican	1	0.42
Bolivian, Canadian	1	0.42
British	80	25.21
British/Brazilian	1	21.00
British/French	1	27.42
British/Irish	1	20.00
British/Nigerian	1	20.00
Colombian/Italian	1	6.75
Dominican	11	0.50
Egyptian	1	0.50
French	1	5.25
Greek	2	2.46
Hungarian, British	1	8.66
Irish	2	6.25
Italian	2	2.21
Jordanian	1	0.33
Nigerian	3	3.06
Norwegian	1	1.16
Polish	2	3.00
Portuguese	3	0.45
Romanian	3	1.22
Spanish	1	0.50
Swedish	1	0.50
Welsh	1	21.00
Zimbabwean	1	2.42
<i>Grand Total</i>	215	12.06

*N.B. Time spent in UK presented in years*

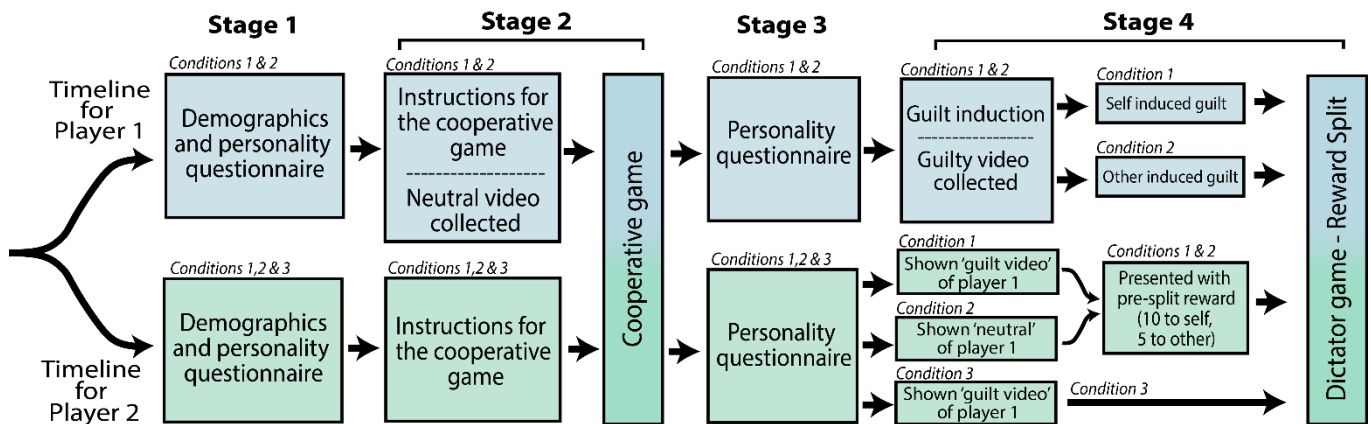


One hundred and twenty-eight participants had European nationality (81 female; mean age = 25.30,  $SD = 8.35$ ) and 87 participants had Asian nationality (including some dual ethnicities; 55 female; mean age = 23.44,  $SD = 5.46$ ; see Table 6.1). Cultural identity was established by self-identification: we asked participants to indicate the ethnical group they identified as. We created a new variable, “cultural congruency”, for further analysis. If the two players shared the same cultural background, they each received a score of 1; if they did not, they both received a score of 0. We included this variable in our models to look at the impact of shared ethnicity on the outcomes of guilt.

All participants were compensated £10 for their time. The experiment lasted for approximately 50 minutes. The projects have been reviewed and approved by the Science Faculty Ethics Committee (SFEC) from the University of Portsmouth (Appendix 5).

### **Experimental design**

Participants acting as Player 1 took part in a between-subject 2 (culture) x 2 (blamed vs not blamed) study. Participants acting as Player 2 took part in a between-subject 2 (culture) x 2 (guilty video vs control video) study. The dependent variable was the division of monetary reward. For Player 2 there was an additional control condition designed as a baseline to check the role of a potential guilt expression in the absence of context (see Figure 6.1).



**Figure 6.1. General Design.**

A flowchart representing the procedure of the experiment.

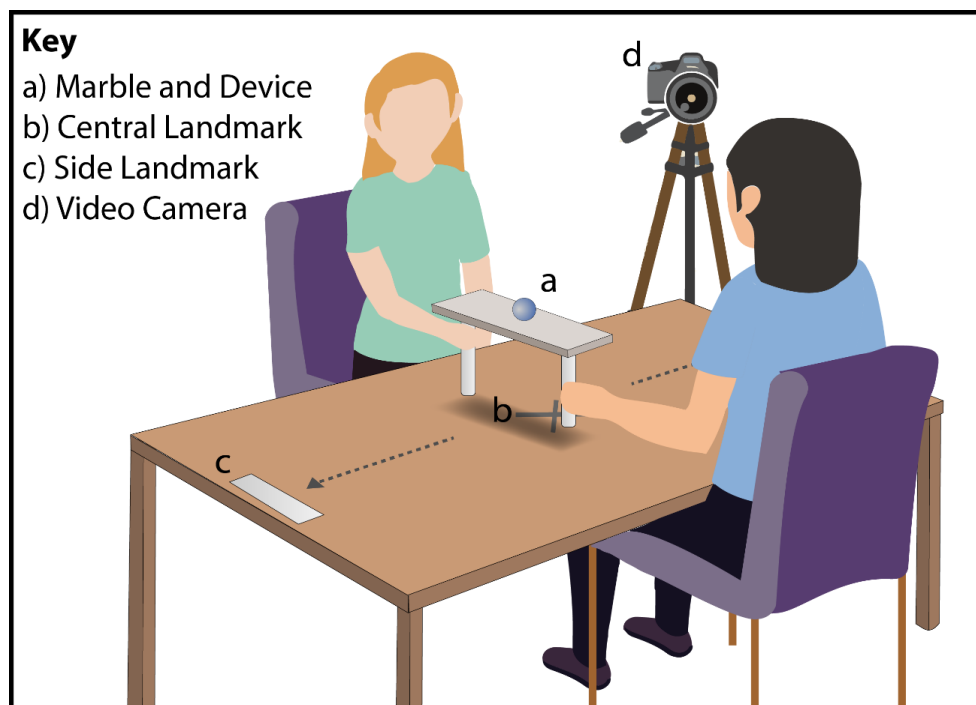
## Procedure

Participants were told that the purpose of the study was to examine how people work together as part of a team. Each dyad began the study in the same room (strangers were introduced to each other) but was then separated. Participants were randomly attributed to a role (Player 1 or Player 2; see Figure 6.1). In **Stage 1**, participants were separated to complete questionnaires (GASP - Cohen et al., 2011; the Unidimensional Relationship Closeness Scale, URCS - Dibble, Levine, & Park, 2012; "How am I in general" - John, Naumann, & Soto, 2008; and the mood check PANAS - Watson et al., 1988; see Appendix 1 for details). Question order was randomised between participants. The experimenter then explained the cooperative task individually to each participant. The cooperative task (see Figure 6.2) was then explained and a video taken of Player 1 to provide stimuli for Player 2 (Stage 4).

Participants were then reunited for **Stage 2**: a cooperative game where the participants had to work together to move a device from one end of the table to the other without dropping a marble (inspired from Miossec & Kheddar, 2009; see Figure 6.2). Participants were told that at the beginning of the game, they were attributed to a shared reward of £20. However, they were informed that the shared reward would reduce each time the marble fell. Participants were also informed that individual cooperative abilities and fine motor function were assessed during the

task in order to determine who was responsible for the majority of marble drops. On completion of the task, the participants were separated again for **Stage 3**, where they completed more questionnaires (mood check, a friendship closeness scale, and the Dirty Dozen - DD; Jonason & Webster, 2010). In **Stage 4** both participants were then randomly allocated into different conditions (see below) and completed one final mood check before being asked to split the reward between their partner and themselves.

At the end of the study, all participants were debriefed and informed of the real goal of the study. They were also reassured that their performances during the game were not assessed and that the feedback provided was fake.



**Figure 6.2. Cooperative game.**

For the cooperative game, two individuals were required to balance a marble on top of a single device that required two people to operate. They were instructed that the aim of the task was to see how many times the device (with balancing marble) could be moved from one end of a table to another, without the marble falling within three minutes. If/when the marble falls – the participants would repeat the task again until the time was up.

**Stage 4 – Player 1.** Player 1 was informed that his/her individual performance on the task was responsible for their failure as a team and that the shared reward would be £15 instead of £20. The experimenter explained that we were monitoring concentration and motor abilities through observational analysis and that his/her concentration and motor coordination was lower than their partner's. During this feedback session, the face and upper body of Player 1 were video-recorded to provide stimuli for Player 2. The player was asked to split the reward (£15) between themselves and their partner (i.e. a dictator game, Forsythe et al., 1994). This money split was designed as a measure of relationship repair, potentially to mend the previous transgression. Player 1 was then randomly allocated into a 'blamed' condition (told that Player 2 understands that it is their fault that they failed and thinks it would be unfair to split the money equally: other-induced guilt) or not blamed (told nothing: self-induced guilt). Due to financial limitation, and to keep the design as simple as possible, we did not manipulate whether Player 1 had reason to feel guilty. This allowed us to minimise the number of variables, already extensive, and reduce the number of participants needed.

**Stage 4 – Player 2.** Player 2 was allocated to one of three conditions. In condition 1, they were told that the £20 reward was reduced to £15 due to the poor performance of Player 1, but that Player 1 had already split the £15 reward in Player 1's favour (10 for P1/5 for P2). This unfair original split was designed as a social transgression and an injustice. We wanted to measure Player 2's propensity to rectify a social injustice by assessing the change made to this original split. They were then presented with a 'guilty' video of Player 1 (when Player 1 received the poor feedback from the game) and were offered the chance to change the reward split (dictator game, Forsythe et al., 1994). Condition 2 was identical but the video of Player 1 was 'neutral' (video taken when the game was explained in stage 1). An additional control condition (Condition 3) included the 'guilty' video but in the absence of any contextual information (not told that it was Player 1's fault and not told that Player 1 had pre-split the reward; both

participants shared the failure). In all conditions, Player 2 was asked to judge the emotional state of Player 1 (including how guilty they thought Player 1 looked in the video).

## Measures

**Facial expression.** The stimulus videos of Player 1 were coded using the Facial Action Coding System (FACS; Ekman & Friesen, 1978; Ekman et al., 2002), using the Interact© software (Mangold, 1998). To ensure unbiased results, reliability analysis was conducted with another certified FACS coder on 22 videos out of the 215 video clips (~10%) extracted from the study videos. The agreement calculation (Wexler's agreement) was calculated based on Ekman, Friesen, and Hager's Investigator's Guide (Ekman et al., 2002). Agreement between coders was 72%, considered a good agreement in FACS methodology (Ekman et al., 2002) and considering the coding scheme used in the present study.

To reduce the FACS dataset to try and identify common and meaningful facial movements, we used a binomial exact test as criteria for exclusion - if any AU/AD was produced by fewer players than the calculated criteria, this AU/AD was removed from the dataset. The binomial exact test allowed us to keep facial movements produced significantly more than chance: if at least 62 participants produced the movement, then this movements reliably occurs across most participants and is not resulting from individual differences. We also ran a correlation between each AU and guilt (self-reported or judged) in order to select AUs that were produced more in association with a higher level of guilt. In order to look at facial signals associated with felt guilt, we looked at all the guilty videos collected from Player 1 to assess any facial movements associated with felt guilt (i.e., production). We ran a correlation between each AU and the self-reported guilt; this left us with 20AU/ADs (out of a possible 41 observed in our data; see *Table 6.2*). In order to explore facial signals perceived as guilt by Player 2, we looked at all the videos presented to Player 2 (i.e., a mix of control

and guilt videos), to assess any facial movements associated with perceived guilt (i.e., perception). We ran a correlation between each AU and the judged guilt; this left us with 12AU/ADs for perception (out of a possible 41 observed in our data; see Table 6.2).

**Table 6.2. List of Action Units (AUs), movements, and miscellaneous actions coded for**

AU	AU Description	Group	Production	Perception
1	Inner Brow Raiser	Upper Face AU		
2	Outer Brow Raiser	Upper Face AU		
4	Brow Lowerer	Upper Face AU		
5	Upper Lid Raiser	Upper Face AU		
6	Cheek Raiser	Upper Face AU		
7	Lids Tight	Upper Face AU		
9	Nose Wrinkle	Lower Face AU		
10	Upper Lip Raiser	Lower Face AU		
11	Nasolabial Furrow Deepener	Lower Face AU		
12	Lip Corner Puller	Lower Face AU		
14	Dimpler	Lower Face AU		
15	Lip Corner Depressor	Lower Face AU		
16	Lower Lip Depressor	Lower Face AU		
17	Chin Raiser	Lower Face AU		
18	Lip Pucker	Lower Face AU		
20	Lip Strech	Lower Face AU		
22	Lip Funneler	Lower Face AU		
23	Lip Tightener	Lower Face AU		
24	Lip Presser	Lower Face AU		
28	Lips Suck	Lower Face AU		
25	Lips Part	Lip Parting		
26	Jaw Drop	Jaw Opening		
27	Mouth Strech	Jaw Opening		
51	Turn Left	Head Position		
52	Turn Right	Head Position		
53	Head Up	Head Position		
54	Head Down	Head Position		
55	Tilt Left	Head Position		
56	Tilt Right	Head Position		
57	Head Forward	Head Position		
58	Head Backward	Head Position		
M59	Nodding	Head Position		
61	Eyes Left	Eye Position		
62	Eyes Right	Eye Position		
63	Eyes Up	Eye Position		
64	Eyes Down	Eye Position		
37	Lip Wipe	Miscellaneous		
50	Talking	Miscellaneous		
Laugh	Laughing	Miscellaneous		
Face Touch	Touching any part of face with hand	Miscellaneous		
Neck Touch	Touching neck with hand	Miscellaneous		
Hair Touch	Touching hair with hand	Miscellaneous		
Ear Touch	Touching ear with hand	Miscellaneous		
Scratching	Any form of scratching on face/near/t	Miscellaneous		

Note: AU10 and AU11, as well as AU23 and AU24 and AU26 and AU27, were combined for analysis

**Dictator game - reward split**

In the dictator game presented to Player 1 (Figure 6.1), players were forced to make an uneven split and attribute at least £1 more to themselves or to their partner. Therefore the dependent variable was the overall proportion of coins given to Player 2 (of the 15 available: *percentage* < 0.5: more money retained for self; *percentage* > 0.5: more money to Player 2).

In conditions 1 and 2, Player 2 was presented with the reward money already split (10 to Player 1, 5 to Player 2). Here, the dependent variable was the magnitude of change from the original split to the final split:  $\frac{10-x}{15}$ , with  $x$  being the final number of coins given to Player 1. In condition 3, Player 2 was presented with an un-split reward and the dependent variable was calculated as for Player 1.

**Demographic and personality data**

We asked each participant to fill out a demographic questionnaire, providing information regarding the ethnic group they self-identified with, the number of time spent in the UK, and their legal nationality. They also filled out multiple personality questionnaires and mood checks at different times of the study. Overall, they completed the Guilt and Shame Proneness scale (GASP; Cohen et al., 2011), the “How am I in general” (John et al., 2008) and the Dirty Dozen (Jonason & Webster, 2010). Additionally, both players completed two measures of friendship: one before (Unidimensional Relationship Closeness Scale - URCS; Dibble et al., 2012) and one after (a friendship closeness scale: “After playing together, how close do you feel to your partner?”, using a 0-100% scale) playing the cooperative game. Finally, each player completed twice the Positive and Negative Affect Scales (PANAS; Watson et al., 1988. For details on all questionnaires, see Appendix 1).

### **Statistical Analysis**

To identify any facial movements associated with guilt (felt or perceived), we ran a factor analysis (Principal Component Analysis with Varimax rotation, PCA) on the frequency data of the 20 (production) or 12 (perception) AU/ADs. PCA was conducted on frame-by-frame data to produce factors consisting of temporally clustered AU/ADs, as previously done by Stratou et al. (2017). As previously shown, conducting a frame-by-frame PCA on FACS data increases the link between cognitive processes and resulting factors of facial movements (Stratou et al., 2017). We added the self-reported guilt in the PANAS (0-5, not guilty to very guilty; self-reported by Player 1) or judged guilt (judged by Player 2) to the corresponding PCA in order to identify movements that were produced frequently more often when guilt was reported or judged. We kept the factor containing AUs clustering with either the reported guilt or the judged guilt for the following analysis. PCA was conducted in SPSS version 24 (IBM, 2016).

Second, we investigated the relationship between the facial movements, guilt, and the Dictator game responses (reward split) using a generalized linear model analysis approach (GLM). Looking at Player 1's response, PCA factor clustering with felt guilt, felt guilt, friendship (median split: 0 acquaintances or 1 close friend), condition, and cultural congruency (whether their reported cultural identity matched that of the other player: EU-EA, EU-EU, EA-EA, and EA-EU) were set as our predictor variables, and the reward split as our dependent variable. Looking at Player 2's response, PCA factor clustering with judged guilt, condition (1: guilt video or 2: control video), cultural congruency (EU-EA, EU-EU, EA-EA, and EA-EU), and the interaction friendship (median split: 0 acquaintances or 1 close friends) x judged guilt were set as our predictor variables, and the reward split as our dependent variable. We ran one last GLM on Player 2's response in the control condition (Condition 3, guilt video, shared failure) with friendship (0 or 1), and cultural congruency (EU-EA, EU-EU,



EA-EA, and EA-EU) set as our predictor variables, and the reward split as our dependent variable.

Finally, we looked at the impact of various personality traits, as well as perceived responsibility, on reported guilt (self-reported or judged) and reward split using GLMs. The GASP scales, the Big Five items, and the annoyance felt toward self and partner were set as our predictor variables in three different models, and the reported guilt (self-reported or judged) was set as our dependent variable. We fitted the GLMs using the function `glm` provided by the packages `lme4` and `lmerTest` for RStudio (Bates et al., 2014). To assess for successful induction of guilt in Player 1, we compared the data collected through the PANAS questionnaires (before vs. after game) using a within-subjects *t*-test. We tested for a change in positive and negative affect before vs. after induction, and additionally, some specific emotional changes in guilt, shame, distress and pride (see Appendix 1 for details).

### 6.3. Results

#### Feeling guilty (Player 1)

##### Self-reported guilt

After receiving the feedback in Stage 4, Player 1 experienced an increase in negative affect ( $M = -0.2337$ ,  $CI = [-0.318; -0.150]$ ;  $t(106) = -5.512$ ,  $p < 0.001$ ,  $d = 0.35$ ), and decrease in positive affect ( $M = 0.4233$ ,  $CI = [0.310; 0.537]$ ;  $t(106) = 7.397$ ,  $p < 0.001$ ,  $d = -0.47$ ). More specifically, participants reported increased guilty feelings ( $M = -0.664$ ,  $CI = [-0.897; -0.430]$ ;  $t(106) = -5.624$ ,  $p < 0.001$ ,  $d = 0.56$ ) and shame feelings ( $M = -0.533$ ,  $CI = [-0.721; -0.344]$ ;  $t(106) = -5.6$ ,  $p < 0.001$ ,  $d = 0.53$ ), and decreased pride ( $M = 0.523$ ,  $CI = [0.321; 0.726]$ ;  $t(106) = 5.116$ ,  $p < 0.001$ ,  $d = -0.44$ ). No changes were found in self-reported distress ( $M = -0.140$ ,  $CI = [-0.295; 0.0143]$ ;  $t(106) = -1.799$ ,  $p = 0.075$ ,  $d = 0.13$ ). Players reported a significantly

higher level of guilt than shame after the induction task ( $M = -0.187$ ,  $CI = [-0.373; -0.0005]$ ;  $t(106) = -1.988$ ,  $p = 0.0494$ ,  $d = 0.17$ ), suggesting that guilt was the primary induced emotion (means and SE presented characterise the difference between the values before T2 and the values after T3 the game; see Table 6.3). The level of friendship did not affect how Player 1 felt after receiving the feedback ( $\beta = 0.042$ ;  $SE = 0.068$ ;  $p = 0.536$ ).

**Table 6.3. Mean and standard deviation of the self-reported emotions throughout the study**

<i>Emotion</i>	<i>Participant</i>	<i>Measure</i>	<i>T1</i>	<i>T2</i>	<i>T3</i>
Guilt	Participant 1	Mean	1.262	1.533	2.196
		SD	0.731	0.883	1.328
	Participant 2	Mean	1.403	1.722	1.347
		SD	0.899	0.923	0.790
Shame	Participant 1	Mean	1.252	1.477	2.009
		SD	0.646	0.781	1.120
	Participant 2	Mean	1.319	1.542	1.310
		SD	0.901	0.903	0.623
Distress	Participant 1	Mean	1.579	1.710	1.850
		SD	0.858	1.064	1.053
	Participant 2	Mean	1.764	1.667	1.479
		SD	1.157	1.007	0.892
Pride	Participant 1	Mean	2.271	2.570	2.047
		SD	1.225	1.198	1.119
	Participant 2	Mean	2.375	2.389	2.704
		SD	1.131	1.205	1.269

*T1: beginning of the study; T2: after the game; T3: after feedback; Analysis conducted between T2 and T3.*

### **Felt guilt and facial movements**

The first PCA produced an 8-factor solution. Felt guilt clustered with AU28 (Lips Suck) and AU37 (Lip Wipe) in the fourth factor, which explained 5.90% of the variance. This factor was retained for further analysis.

**Reward split**

On the whole Player 1 gave a bigger share to Player 2, keeping an average of 6 coins (out of 15) for themselves ( $M = 5.972$ ,  $SD = 1.88$ ) and giving 9 coins to their partner ( $M = 9.03$ ,  $SD = 1.88$ ;  $t(106) = -8.405$ ,  $p < 0.001$ ). Player 1 gave more money to their partner the guiltier they felt ( $\beta = 0.022$ ;  $SE = 0.0096$ ;  $p = 0.0262$ ,  $r^2 = 0.13$ ). European Player 1 also gave less money to their partner of Asian cultural background ( $\beta = -0.12$ ;  $SE = 0.055$ ;  $p = 0.0257$ ) and gave less money when producing increased rates of Factor 1B (AU22 and Face Touch;  $\beta = -0.0392$ ;  $SE = 0.0173$ ;  $p = 0.0253$ ). There was no effect of friendship on reward split ( $\beta = 0.0040$ ;  $SE = 0.0063$ ;  $p = 0.523$ ). Further analysis showed that the strength of the friendship was not influenced by the cultural congruency between players ( $\beta = 0.338$ ;  $SE = 0.586$ ;  $p = 0.565$ ). Whether Player 1 was blamed or not for his/her performance did not influence reward split ( $\beta = 0.00331$ ;  $SE = 0.239$ ;  $p = 0.890$ ).

We investigated the relationship between the GASP scales, the Big Five items, annoyance, and the reward split. Two subscales of the GASP scale correlated positively with Player 1's decision (see Sup Mat for details on GASP): one of the guilt subscales – Negative Behaviour-Evaluations ( $\beta = 0.0225$ ;  $SE = 0.0100$ ;  $p = 0.027$ ), and one of the shame subscales – Negative Self-Evaluations ( $\beta = 0.0267$ ;  $SE = 0.0114$ ;  $p = 0.0204$ ). Moreover, we found a positive relationship between money split and Conscientiousness: the higher players rated on this item, the more money they gave their partner ( $\beta = 0.0499$ ,  $SE = 0.0204$ ;  $p = 0.0164$ ). The level of annoyance felt toward oneself also affected the final reward split: the more annoyed players reported being at themselves, the more money they gave their partner ( $\beta = 0.00082$ ,  $SE = 0.000399$ ;  $p = 0.0423$ ).

### Personality

We investigated the relationship between the GASP scales, the Big Five items, and annoyance, and the reported guilt. All four subscales of the GASP scale correlated positively with the reported guilt: both guilt subscales – Repair action tendency ( $\beta = 0.425$ ;  $SE = 0.159$ ;  $p < 0.01$ ), and Negative Behaviour-Evaluations ( $\beta = 0.267$ ;  $SE = 0.105$ ;  $p = 0.0128$ ); and both shame subscales – Withdrawal action tendencies ( $\beta = 0.289$ ;  $SE = 0.108$ ;  $p < 0.01$ ), and Negative Self-Evaluations ( $\beta = 0.346$ ;  $SE = 0.119$ ;  $p < 0.01$ ). Moreover, we found a positive relationship with both Agreeableness ( $\beta = 0.695$ ;  $SE = 0.283$ ;  $p = 0.160$ ) and Emotional stability ( $\beta = 0.440$ ;  $SE = 0.201$ ;  $p = 0.308$ ). The level of annoyance felt toward oneself (see Table 6.4) affected greatly the level of felt guilt reported: the more annoyed they were at their performance during the game, the guiltier they felt after ( $\beta = 12.328$ ;  $SE = 1.832$ ;  $p < 0.0001$ ).

**Table 6.4. Annoyance towards self and partner measured after the cooperative game**

<i>Emotion</i>	<i>Participant</i>	<i>Measure</i>	
Annoyance towards self	Participant 1	Mean	40.632
		SD	30.429
	Participant 2	Mean	9.800
		SD	16.369
Annoyance towards partner	Participant 1	Mean	19.000
		SD	23.465
	Participant 2	Mean	10.260
		SD	18.621

## Perceiving guilt in others (Player 2)

### Manipulation check – felt guilt

After receiving the feedback that Player 1 was at fault during the cooperative task failure (conditions 1 and 2), Player 2 experienced a decrease in negative affect ( $M = 0.156$ ,  $CI = [0.0464; 0.265]$ ;  $t(71) = 2.842$ ,  $p < 0.01$ ,  $d = -0.28$ ), and also a decrease in positive affect ( $M = 0.283$ ,  $CI = [0.110; 0.456]$ ;  $t(71) = 3.263$ ,  $p < 0.01$ ,  $d = -0.28$ ). There was also a significant decrease in guilty feelings ( $M = 0.375$ ,  $CI = [0.150; 0.600]$ ;  $t(71) = 3.329$ ,  $p < 0.01$ ,  $d = -0.43$ ), and shame feelings ( $M = 0.239$ ,  $CI = [0.0415; 0.437]$ ;  $t(70) = 2.413$ ,  $p = 0.0184$ ,  $d = -0.25$ ), and a significant increase in pride ( $M = -0.338$ ,  $CI = [-0.621; -0.0553]$ ;  $t(70) = -2.384$ ,  $p = 0.0198$ ,  $d = 0.26$ ). No changes were found in self-reported distress ( $M = 0.197$ ,  $CI = [-0.0131; 0.407]$ ;  $t(70) = 1.870$ ,  $p = 0.0656$ ,  $d = -0.19$ ). The level of reported guilt after receiving the feedback did not differ significantly from the level of reported shame ( $M = -0.0563$ ,  $CI = [-0.211; 0.0980]$ ;  $t(70) = -0.728$ ,  $p = 0.469$ ,  $d = -0.05$ ; *means and SE presented characterise the difference between the values before and the values after the game*).

In the control condition (condition 3), after receiving the feedback, Player 2 experienced a decrease in negative affect ( $M = 0.119$ ,  $CI = [0.0003; 0.239]$ ;  $t(35) = 2.036$ ,  $p = 0.049$ ), a decrease in positive affect ( $M = 0.142$ ,  $CI = [0.0038; 0.279]$ ;  $t(35) = 2.087$ ,  $p = 0.044$ ) and a significant decrease in shameful feelings ( $M = 0.361$ ,  $CI = [0.0678; 0.654]$ ;  $t(35) = 2.499$ ,  $p = 0.0172$ ). There were no differences in the levels of guilt, distress or pride reported after receiving the feedback.

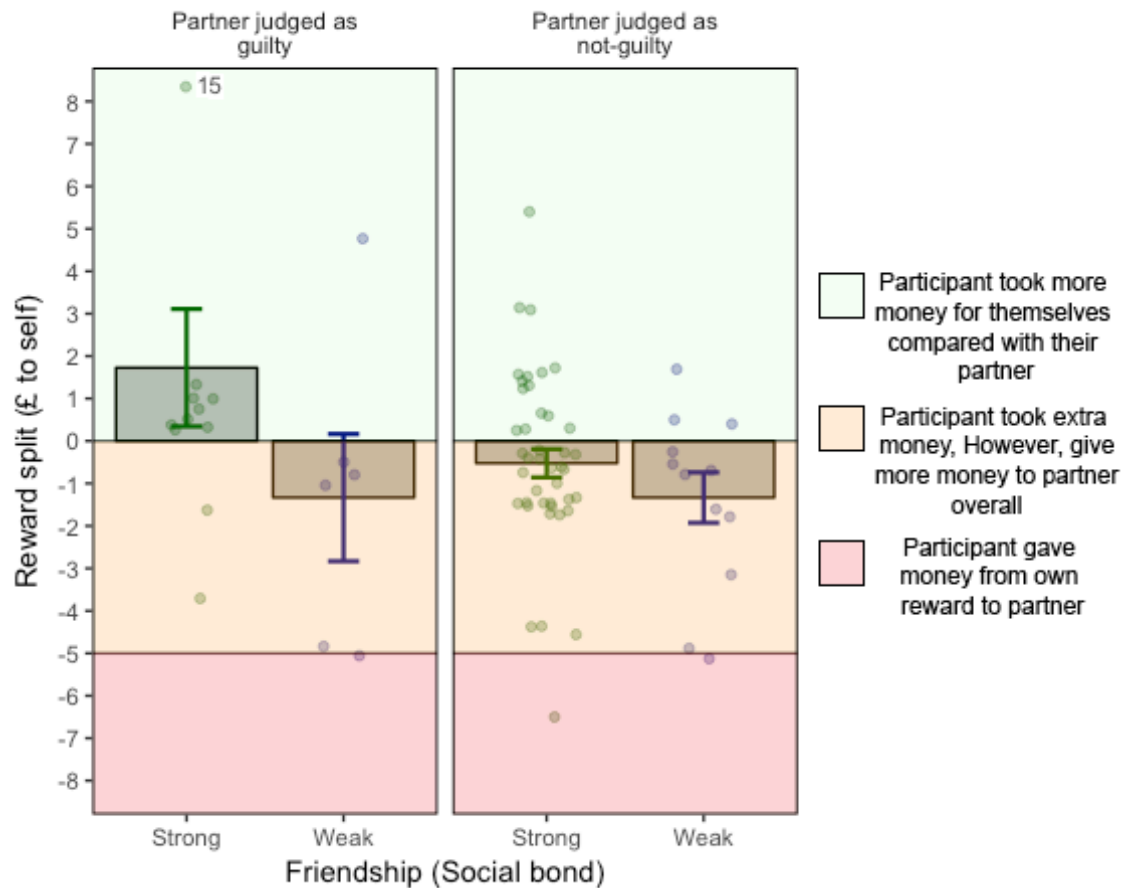
### Judged guilt and facial movements

We analysed the stimulus videos in relation to judged guilt from Player 2, to see if there were any patterns of movements identified as guilt: the PCA produced a 4-factor solution. Judged guilt clustered on its own in the fourth factor explaining 7.81% of the variance and did not cluster with any AUs/ADs.

We investigated which factors influenced the judgement of guilt by Player 2: condition, cultural congruency and friendship. None of the investigated variables affected judged guilt: condition  $\beta = -0.13$ ,  $SE = 0.12$ ,  $p = 0.286$ ; cultural congruency  $\beta = [0.25; 0.84]$ ,  $SE = [0.22; 0.45]$ ,  $p > 0.685$ ; friendship index  $\beta = 0.028$ ,  $SE = 0.058$ ,  $p = 0.632$ ,  $r^2 = 0.083$ . The self-reported guilt also had no effect on Player 2's judgement of guilt in their partner ( $\beta = 0.0530$ ;  $SE = 0.0751$ ;  $p = 0.482$ ,  $r^2 = 0.050$ ).

### **Reward split**

On the whole, Player 2 did not favour themselves over their partner in the reward split, keeping 7 coins (out of 15) on average for themselves ( $M = 7.39$ ,  $SD = 1.46$ ) and giving 8 coins to their partner ( $M = 7.69$ ,  $SD = 1.44$ ;  $t(106) = -0.870$ ,  $p = 0.387$ ). However, this does mean that Player 2 changed the original split presented to them, from 10 coins for their partner to 8 on average ( $t(106) = 13.515$ ,  $p < 0.001$ ). Player 2 was more likely to change the original split (i.e., taking more money for themselves than the original 5 coins) the more guilty they judged Player 1 to feel ( $\beta = 0.031$ ;  $SE = 0.011$ ;  $p < 0.01$ ,  $r^2 = 20$ ). In addition, friendship interacted with the judgement of guilt to impact Player 2's decision: when Player 1 was judged as looking not-guilty (judged guilt  $< 3$ ), Player 2 changed the original split but still attributed more coins to Player 1, regardless of their friendship (Figure 6.3, right side); however, when Player 1 was judged as looking guilty (judged guilt  $\geq 3$ ), Player 2 claimed more money when playing with a friend (index  $>$  median 3.67) ( $\beta = 0.0810$ ;  $SE = 0.0305$ ;  $p < 0.01$ ; Figure 6.3, left side), attributing more money to themselves than to their partner. There was no effect of cultural congruency ( $\beta > 0.0124$ ;  $SE [0.0231; 0.0673]$ ;  $p > 0.402$ ).



**Figure 6.3. The interaction between judged guilt and friendship.**

Reward split is presented as a difference between the final amount taken by P2 and the amount attributed to P1. Friends were split into “Weak friendship” and “Strong friendship” based on a median split: if friend index was higher than 3.67, participants were considered as close friends. Judged guilt was turned into a binary variable for this graph, using a median split: if judged guilt was higher than 3 on the PANAS, Player 2 judged Player 1 as guilty (‘Guilty’); if the judged guilt was lower than 3, Player 2 did not judged Player 1 as guilty (‘Not Guilty’). The line at -5 represent the original split: 5 coins attributed to P2 and 10 coins attributed to P1. The line at 0 represents financial equality (which could not be obtained as we presented Player 2 with an odd number of coins). Anything above the middle line means Player 2 claimed more money for themselves.

### Personality

**Felt guilt.** We investigated the relationship between the GASP scales, the Big Five items, and annoyance, and the self-reported guilt, looking at all conditions together. Two subscales of the GASP scale correlated positively with Player 2 felt guilt: both shame subscales – Withdrawal action tendencies ( $\beta = 0.162$ ; SE = 0.0721;  $p < 0.01$ ), and Negative Self-Evaluations ( $\beta = 0.169$ ; SE = 0.0689;  $p < 0.01$ ). The level of annoyance felt toward oneself

affected greatly the level of felt guilt reported: the more annoyed they were at their performance during the game, the guiltier they felt after ( $\beta = 0.00175$ ;  $SE = 0.00481$ ;  $p = 0.0372$ ). Moreover, we found that the level of annoyance felt toward oneself affected the level of judged guilt: the more annoyed they were at their own performance during the game, the guiltier they thought Player 1 was afterwards ( $\beta = 0.00133$ ;  $SE = 0.00602$ ;  $p = 0.0292$ ).

**Judged guilt.** We investigated the relationship between the GASP scales, the Big Five items, annoyance towards self or partner, and the judged guilt, looking at all conditions together. None of the personality traits of Player 2 influenced the way they judged Player 1's guilt (GASP:  $p > 0.140$ ; Big Five:  $p > 0.0821$ ). However, we found that level of annoyance felt toward oneself influenced the level of judged guilt: the more annoyed they were at their own performance during the game, the guiltier they thought Player 1 looked in the video ( $\beta = 0.0133$ ;  $SE = 0.00602$ ;  $p = 0.0292$ ). The annoyance felt towards their partner did not influence the judged guilt ( $\beta = -0.00852$ ;  $SE = 0.00763$ ;  $p = 0.267$ ).

**Reward split.** We investigated the relationship between the GASP scales, the Big Five items, annoyance at their partner, and the final split, looking at all conditions together. One subscale of the GASP scale related negatively with Player 2's final decision: the guilt Repair action tendencies subscale ( $\beta = -0.0269$ ;  $SE = 0.0117$ ;  $p = 0.0228$ ). None of the other personality data collected influenced the final decision: Big Five index,  $p > 0.101$ ; annoyance at partner,  $\beta = -0.000524$ ;  $SE = 0.000548$ ;  $p = 0.341$ ; annoyance at self,  $\beta = -8.07e-05$ ;  $SE = 4.33e-04$ ;  $p = 0.853$ .

**Absence of context (control condition 3).** Player 2's response in the control condition was examined in a GLM with friendship and cultural congruency set as our predictor variables, and the reward split as our dependent variable. Player 2 tended to give more money to their partner, keeping 7 coins on average for themselves ( $M = 7.08$ ,  $SD = 1.34$ ) and giving 8 coins to their partner ( $M = 7.92$ ,  $SD = 1.44$ ;  $t(35) = -1.875$ ,  $p = 0.0692$ ). Player 2's reward split was not affected by judged guilt ( $\beta = -0.00734$ ,  $SE = 0.0189$ ,  $p = 0.700$ ), or



friendship ( $\beta = 0.00101$ ,  $SE = 0.00819$ ,  $p = 0.903$ ), but European Player 2 gave less money to European Player 1 ( $\beta = -0.077$ ,  $SE = 0.036$ ,  $p = 0.0402$ ).

#### 6.4. Discussion

In the present experiment, we examined the social consequences of both feeling guilty and observing guilt in others. First, we examined the behavioural outcomes associated with guilt and the facial movements associated with guilty feelings. Second, we examined the behavioural outcomes of being the victim of wrongdoing, and how this is affected by friendship. We found that participants who were told they were responsible for the failure of the cooperative game, allocated a greater proportion of the reward to their partner the more guilty they felt (regardless of the strength of their friendship). Similarly, players punished partners more when they looked guilty, and this effect was modulated by friendship: players punished their close friends more when they were perceived as feeling guilty.

When told they performed poorly, participants gave more money to their partners the guiltier they felt, regardless of their friendship, which is consistent with previous findings highlighting the positive social consequences associated with feelings of guilt (J. Carroll, 1985; Cryder et al., 2012; Ketelaar & Au, 2003; Tangney, 1999). This could function as both reputation and relationship repair (Haley & Fessler, 2005); acknowledging that the wrongdoing was committed, but also indicating that the wrongdoing was unintentional and further retaliation is not necessary (Henrich & Boyd, 2001; Jensen, 2010). The participants' personality traits affected some of the decisions made. Pro-social personality traits (Agreeableness and Emotional stability) correlated positively with the level of reported guilt. This supports the idea that guilt can be adaptive and have a positive social function (Baumeister et al., 1994; Cohen, Panter, & Turan, 2012; Tangney et al., 2007). Moreover, conscientious Player 1 gave a bigger share to Player 2;

this personality trait is linked to diligence and could indicate a propensity to doing the right thing and making amend. The annoyance felt towards oneself also had an interesting effect on the decisions made by both players. On the one hand, Player 1 reporting higher self-annoyance felt guiltier and gave more money to Player 2. On the other hand, Player 2 reporting a higher level of self-annoyance judged Player 1 as looking guiltier than Player 2 reporting low self-annoyance. Player 2 might have projected their own emotions onto Player 1, assuming the official wrongdoer would feel as bad as they did.

We also found a small effect of culture, with European wrongdoers allocating a lesser share of the reward to their Asian partners. There is no clear explanation for this result; all our participants were tested in the U.K., where our Asian participants decided to live, which might change their in-group/out-group vision. We would argue that it is not unexpected to have found an in-group bias (Tajfel, 1978; Tajfel, Turner, Austin, & Worchel, 1979) in the European wrongdoers, but it is more surprising to not have found it in the Asian wrongdoers. Testing a larger sample would help clarify this result. Being blamed by the partner did not add to the existing guilt. This supports previous findings by Parkinson and Illingworth (2009) of a slight tendency for blame to reduce guilt in the presence of self-blame.

When presented with an unfair reward split, the victim changed this unequal split to a fairer split (i.e., took more money for themselves than the original £5) and did so more when they judged the wrongdoer to be guiltier. This is in contrast with previous studies showing people being more lenient towards people expressing guilt (Bandes, 2016; De Jong et al., 2003; Weisman, 2014). One reason might be that our study examines the function of guilt between interacting participants, rather than via a hypothetical game or autobiographical recall. Alternatively, it could be that participants evaluate the contribution of their partner differently in light of the guilty look. The guilty look might reinforce the information given to the participant that their partner has performed poorly and act as an acknowledgement of the wrongdoer's bad

behaviour (Rosenstock & O'Connor, 2018). Moreover, this punishment function of guilt was further enhanced when victims had a strong friendship with the wrongdoer suggesting that the victim punished their guilty friends more. People sharing a strong friendship bond have invested a lot of time and resources into the relationship; any disruption to the relationship ought to be repaired in order to save and restore the relationship (Jensen, 2010), regardless of short-term consequences such as punishment. Finally, Player 2 that scored higher on the repair action tendency subscale of the GASP questionnaire attributed more money to themselves than to their partner. The repair action tendency subscale measure the propension someone has to engage in reconciliatory or pro-social behaviours after committing a wrongdoing. It could be understandable that someone prone to make amend expect others to do so and be more prone to punish their social partner. This finding supports the idea that guilt has a much stronger function and impact within existing friendships, leading to harsher punishments. This is in line with previous research reporting clearer emotional expressions among friends (Hess et al., 1995; Wagner & Smith, 1991).

We used a bottom-up coding scheme to identify facial patterns associated with the experience (production and perception) of guilt. We found that players informed of their poor performance (Player 1) displayed more lip suck (AU28, Lips Suck) and lip wiping (AU37, Lip Wipe; Ekman & Friesen, 1978) when reporting a higher level of felt guilt. Interestingly, this is different from previously documented facial patterns of guilt (Julle-Danière et al., under review). Such inconsistencies suggest that guilt may not manifest as a uniform pattern of facial movements across guilt-inducing contexts. We also did not identify a pattern of facial movement associated with the judgment of guilt by Player 2. Therefore, we found no direct correspondence between facial movements and perception of guilt, supporting current research that the link between facial expression and secondary emotion could be less strong than previously thought (Barrett et al., 2019). Moreover, we did not find any impact of the face alone on the victim's response (condition 3), and our experimentally induced facial expressions did not have an impact (no

difference between conditions 1 and 2). Therefore, although we found that judgements of guilt have an impact on social outcomes, we do not know which behavioural (or other) factors were influencing these judgements.

### **Conclusion**

This is the first study examining the impact of friendship on the consequences of guilt in social interaction between participants. We examined genuine social interactions between dyads differing in relationship quality and tested the impact of guilt within the interaction. In line with previous research, we demonstrated the pro-social role of guilt: guilty players gave more money to their partners. We also revealed a new finding demonstrating a punishment effect of guilt, and a modulating effect of friendship on victims' behaviours: victims of wrongdoing punished their close guilty friends more than less close guilty friends. It could, therefore, seem costly for the wrongdoer to communicate guilt; however, as punishment after wrongdoing has been demonstrated to benefit the relationship in the long-term (Henrich & Boyd, 2001; Rosenstock & O'Connor, 2018), the long-term benefits could outweigh this cost. The findings demonstrate that the social function of emotions may differ depending on the quality and type of relationship between interacting partners.

## 7. General Discussion

### Overview

The overarching focus of this thesis was to explore what people do with their faces when feeling guilty and investigate the potential communicative function of a facial signal of guilt. First, this thesis provides evidence that people do not produce the same facial signals in guilt-eliciting situations than in neutral social situations. Second, this thesis identified patterns of facial movements people associated with feelings of guilt, showing consistency in the perceived signal across samples of participants. Third, this thesis provides evidence that the perception of guilt, from face or situation, will in turn impact decision making, and that the punishment of a guilty individual could be mediated by friendship and the quality of social relationships. Fourth, this thesis demonstrates the importance of contextual information in situations where secondary emotions arise. The exact nature of the information conveyed by the face might remain debated (see Chapter 1 for review), but this thesis demonstrated that the understanding of facial signals produced in complex social situations is improved by placing faces back in context. The results of this thesis stem from an experimental approach, video coding, and encompassing individuals from diverse ethnic backgrounds; demonstrating the importance of the study of genuine facial expressions. The studies presented in this thesis are concerned with ecological validity, aiming at recreating real social interactions in the lab and using spontaneous, dynamic stimuli. The findings of this thesis provide a stepping-stone for the study of spontaneous displays associated with emotional states, more specifically the ones associated with secondary emotions where context and social environment seemed to affect both the experience and perception of emotions.

## **Main findings and implications**

### **Facial signal of guilt as source of information**

Communication is the cornerstone of successful relationships as a way to share information on facts (share knowledge, agree on future actions) and emotional states with our relatives. This is achieved primarily by the exchange of a set of social signals, using verbal and non-verbal signals, such as facial expressions and body postures (Darwin, 1872; Elfenbein et al., 2007; Marsh et al., 2003). As such, facial displays, the rapid nonverbal transmission of socially relevant information (Shariff & Tracy, 2011), can be used to convey messages regarding internal states, emphasizing or contradicting the verbal speech, or even occurring in its absence (Horstmann, 2003). Facial expressions, more than any other nonverbal signals, are thought to be the physical representations of our emotions (Ekman & Friesen, 1971; Izard, 1997).

In two studies, we aimed to identify facial movements and behavioural displays associated with the experience of guilt in humans (Chapter 2). We examined the production of guilt as well as whether others perceived guilt from the face of those experiencing it. We found a positive relationship between the level of self-reported guilt and the extent this individual was judged as feeling guilty by others. The patterns identified in these experiments showed slight variation between what people do when feeling guilty and what people see when identifying guilt. Moreover, the facial signals associated with guilty feelings were produced more in guilt-eliciting situations. Our results support the idea that guilt could have evolved as an observable phenomenon with a potential communicative social function. In a different study (Chapter 6), we examined genuine social interactions between dyads differing in relationship quality and tested the impact of guilt within the interaction. The data in this thesis were in line with previous research and demonstrated the pro-social role of guilt: guilty players gave more money to their partners. Moreover, the feelings of guilt and the facial signals associated with the experience of guilt were not impacted by the quality of friendship. We also uncovered a new finding

demonstrating a punishment effect of guilt: victims of wrongdoing punished their guilty close friends more than guilty less close friends. It could, therefore, seem costly for the wrongdoer to communicate guilt; however, as punishment after wrongdoing has been demonstrated to benefit the relationship in the long-term (Henrich & Boyd, 2001; Rosenstock & O'Connor, 2018), the long-term benefits could outweigh this cost. It seems thus that the signal of guilt has a communicative social function: the wrongdoer make amends and recognise the wrong done by expressing guilt, perceived by the victim as an acknowledgement of the wrong done and the expression of remorse. However, it should be noted that the facial signals produced by guilty partners did not affect the victim's perception of the wrongdoer's guilt; the general context alone seemed to have influenced the victim.

### **Guilt in context**

In real life, facial expressions always occur in context; however, previous judgement studies have proven that people can reliably identify some facial expressions of primary emotion in the absence of context (Crivelli et al., 2017; Jack et al., 2016; Keltner & Buswell, 1996). Introducing contextual information (i.e., written information, body posture, social environment) to judgment of primary emotions revealed a conflicted picture, revealing the impact of incongruent context on the understanding of the facial expressions presented (Aviezer et al., 2008; J. M. Carroll & Russell, 1996; Fernandez-Dols et al., 1993; Hess et al., 2016; Matsumoto & Sung Hwang, 2010). In study 2b (Chapter 3), we investigated the relative importance of context in interpreting facial signals associated with secondary emotion, but also the universal aspect of perceived guilty signals. We conducted a replication study, presenting videos of guilty individuals with and without contextual information to a sample of US-based students. In this thesis, I demonstrated that individuals from different cultural backgrounds rely on the same facial signals to identify guilt on a face and that facial expression of guilt is better interpreted when the situational context is provided alongside the facial expression. The data presented in this thesis

contribute to the idea that secondary emotions are more idiosyncratic and context-dependent than primary emotions and support the trend for ecologically valid judgment research as a facial display is rarely occurring without context in real life.

### **Guilt and personality**

Context and facial expressions are two extrinsic factors that can influence the perception of guilt on someone's face. An intrinsic factor seems also to play a role in people's perception of guilt: their personality. Throughout the studies presented here, we found that Machiavellianism was positively correlated with the self-reported guilt whereas narcissism was negatively correlated with self-reported guilt (Chapter 2). Moreover, pro-social personality traits (Agreeableness and Emotional stability) correlated positively with the level of reported guilt (Chapter 6), supporting the idea that guilt can be adaptive and have a positive social function (Baumeister et al., 1994; Cohen et al., 2012; Tangney et al., 2007). We also found a correlation between subscales of the GASP scale and the reported guilt. In Chapter 6, we found in both Players 1 and 2 a correlation between both shame subscales (Negative Self-Evaluation and Withdraw) and felt guilt. The Negative Self-Evaluation subscale measures one's propensity to act ethically and the need to be well perceived by your peers (Cohen et al., 2011). However, our studies did not reveal any correlation between personality and judged guilt. Regarding decisions made, conscientious Player 1 gave a bigger share to Player 2; this personality trait is linked to diligence and could indicate a propensity to doing the right thing and making amend (Chapter 6). It is interesting to note that different personality traits affected felt guilt depending on the context the guilt was experienced in.



**Facial expression: emotional signal or predictor of future actions?**

The universality in the perception of basic emotion has also been found to be method-bound (Fridlund, 2017; Gendron, 2017). Specifically, using forced-choice categories seems to amplify ‘correct’ responses and mask the noise with which expressions are spontaneously received (Gendron, 2017; Gendron et al., 2018). Recent studies (Crivelli et al., 2016, 2017) tried to address this issue by using “prototypical” expressions of emotion in a design comparing the Basic Emotion Theory (BET) classical approach (Ekman & Friesen, 1969) to the Behavioural Ecology View (BECV; Fridlund, 2017). Here, the different classification methods led to variation in the accuracy of labelling guilt displays correctly (Chapter 4). Using dynamic, ecologically valid stimuli, we show that people do not reliably identify a facial expression associated with guilt when using explicit emotion or action tendency labels, but do distinguish the expression of guilt from other emotional displays when using dimensional ratings (aroused-asleep; pleasant-unpleasant). Therefore, secondary emotions may be reliably interpreted by observers, but not necessarily using categorical, discrete labels. In other words, secondary emotions might still be associated with reliable signals, but that these are not necessarily conceptualised as discrete phenomena. Moreover, the presence of contextual information would have certainly improved the general accuracy score for guilt displays. Finally, it is possible that the person looking at a guilty face might look for different signals if there are or not directly impacted by the wrong done. In Chapter 2, the judges were not affected by the action of the participants from Study 1, whereas Player 2 in Chapter 6 were directly impacted by Player 1’s actions. This could explain the different results found when looking at facial behaviours associated with judged guilt.

**Why the confusion when looking at guilty faces?**

Perceiving facial expressions and being able to understand their meaning in social interaction is a key element of human interactions. Words might fail to categorise emotional signal, but the way individuals look at faces have been proven to be a reliable way to discriminate

between facial expressions of primary emotions (Beaudry et al., 2014; Blais et al., 2017; Eisenbarth & Alpers, 2011; Vaidya et al., 2014). Moreover, the way faces with 'emotional' content are processed has been shown to differ from how neutral faces are processed (Blais et al., 2008; Kelly et al., 2011; Rodger et al., 2010). Therefore, one way to examine whether facial movements contain information that can be considered 'emotional' is to examine patterns of eye movements and investigate the informative value of the different parts of the face (Jack et al., 2009; Yuki et al., 2007). Here, when looking at facial signals associated with guilt, observers exhibited scanning patterns similar to the ones used when looking at facial displays of primary emotions. Observers did not seem to discriminate between emotional displays by using different scanning patterns, but thanks to the information carried out on the face. It is thus possible that the emotional display of guilt was perceived as a mix of primary emotions rather than as a unique, specific emotion.

## Limitations

The work carried out throughout this thesis remains explorative. It is the first extensive project aiming at identifying a non-verbal signal associated with guilt using a bottom-up experimental approach and trying to understand how this signal is perceived and interpreted by observers. The results presented in this thesis do not provide a definite answer on what the face of guilt is; it provides preliminary results on the production of a non-verbal guilty signal in two specific contexts. The data presented also support the social value of a signal of guilt, as observers were able to identify the presence of guilty feelings using specific non-verbal signals. Even though this thesis is exploratory, I have provided evidence that signals associated with secondary emotions can be studied in the lab, while using ecologically valid methodologies. However, a lot remains to be done to understand the face of guilt and its social functions.

One of the fundamental issues surrounding the study of emotional displays is that it relies on individuals ability to accurately identify how they feel (Barrett, 2017a; Russell & Fernández-Dols, 1997). Being able to identify how someone feels requires as much practice as a designer requires dissociating between various shades of burgundy: it necessitates practice, a good knowledge of oneself, the ability to listen to one's body, and being able to recognise the signs of different emotions. It is nearly impossible for the lambda participant recruited to take part in studies. In the absence of "emotional fingerprints" (Barrett, 2017a; Smith et al., 2018), we are bound to believe that participants honestly report how they think they are feeling at a given time. It is possible that some of the participants misattributed their emotions or reported a lower level of guilt to minimise the negative feelings associated with the realisation that they did something wrong.

Not only did I have to rely on self-report to identify the emotional states of the participants, but also it is possible that the methods used to collect reported feelings of guilt influenced the

results presented in this thesis. Indeed, Tignor and Colvin (2019) recently pointed out that the way emotions, and specifically guilt, were measured could lead to radically opposed findings. Guilt is usually measured using one of two types of questionnaires: situated or unsituated questionnaires. Situated measures assess feelings of guilt by presenting participants with a series of hypothetical transgressions described in written scenarios (e.g., GASP - Tignor&Colvin, 2019; see Appendix 1). In contrast, unsituated measures provide participants with a list of feelings and emotion words and ask them to report how frequently they experience each emotion over a given period of time and across situations (e.g., PANAS – Tignor & Colvin, 2019; see Appendix 1). Situated measures of guilt were found to be positively associated with adaptive interpersonal behaviours, whereas unsituated measures of guilt were associated with neuroticism and characterised maladaptive experiences of guilt (Tignor&Colvin, 2019). However, those unsituated measures, and specifically the PANAS (Watson et al., 1988) used in the studies presented in this thesis, can be adapted for participants to report their emotional state at the present time, turning this unsituated questionnaire into a situated measure of specific emotions (the situation being the context of the experience/the time at which they fill out the questionnaire). By highlighting the relevance of having situated measures, Tignor and Colvin (2019) propose a new model for investigating guilt and evaluating the adaptive function of the feeling based on the current situation. Using this model would confirm the adaptive function of guilt participants reported feeling in our studies (Chapters 2 & 6): participants reported higher levels of guilt in guilt-eliciting situation compared to neutral situations, and when reporting guilt it was in relation to the wrong allegedly committed. It thus seems that the participants in those studies reported adaptive guilt but comparisons between methods and new models would help confirm those results.

Before having to consider the best method to collect accurate and honest emotional states, we had to recruit over nine hundred participants, including three hundred and ninety-three individuals with an Asian cultural background. However, one limitation regarding this PhD work

is the population sampled: the participants were mainly members of the University of Portsmouth (either member of staff or students, mainly non-psychology); some participants were members of the public, local residents from a specific area of the U.K. The main challenge was to recruit individuals from Asia, that had recently arrived in the U.K., speaking and understanding English enough to understand oral instructions (they could use translators whenever they needed to for written instructions), and still being good representatives of their mother-country. I came to realise that on the one-hand, Western individuals had the “research-culture”: it was very easy to recruit Western individuals to take part in studies and experiments; on the other hand, Asian individuals did not have it. It has been challenging finding willing participants coming from Asian countries; for a lot of them, the monetary reward did not seem to be enough of an incentive. However, once I recruited a few individuals, they referred friends and acquaintances. Asian cultures are characterised by close social groups, they are often structured by a sense of duty, where individuals are perceived as part of something bigger (family, company, country; Kim et al., 2008), and people develop strong bonds and cooperate mostly with close relatives (Gächter et al., 2010; Henrich, Ensminger, et al., 2010; Henrich, Heine, et al., 2010). It seems that even though they do not have the “research-culture”, they have a “group-culture” inciting them to collectively take part in an event, even at different moments in time.

Moreover, we need to acknowledge that even though we tried to have as varied a sample as possible, we did so from two geographical points (Portsmouth and San Diego) and one online sample (Chapter 4). The data in this thesis are only suggestive of population differences and are insufficient to claim that participants from different cultural backgrounds present variation in their display or interpretation of guilt. Our results do not allow us to claim for universality of facial expression of guilt, but they do however tend to demonstrate consistency in facial signals used to perceive guilt on a face. Recruiting participants in their home country, while interacting with individuals sharing their culture and nationality might help improve not only the quality of

cross-cultural research and increase variability in our sample, but also the ability to recruit willing participants.

In order to increase the ecological validity of the experiments presented in this thesis, I used spontaneous, genuine, dynamic videos to present emotional stimuli. Doing so might have increased the genuineness of the videos, but what I gained in validity I lost in standardisation. This added another level of complexity that needs to be taken into account in following research using more complex analytical tool, allowing for such complexity in the stimuli (Coco & Dale, 2014; Scherer et al., 2013).

In comparison with previous research (e.g., Beaudry et al., 2014; Schurgin et al., 2014; Crivelli et al, 2016, 2017; Jack et al, 2006), the stimuli presented here were not standardised and presented a lot of variations between them. I wanted to capture genuine facial expressions and as such, I did not ask participants to stay still, facing the camera. At the time of the social interactions, participants were not aware that they were recorded in order to elicit spontaneous facial signals (*N.B. participants were informed at the end of the studies and consent obtained*). Some of the previous findings were not replicated (e.g., Beaudry et al., 2014; Jack et al., 2016; Keltner & Buswell, 1996), probably due to the differences in stimuli; however, we built upon previous work and did replicate many findings that used still, posed images (e.g., Crivelli et al, 2016; Shuman et al, 2017; Gendron, 2017). Because of their imperfections, the stimuli used here are closer to real-life interactions and the results presented in this thesis are a stepping stone in the study of genuine facial expressions.

Overall, it is possible that our results are likewise limited to those specific contexts. Guilt can be induced by various experiences and situations (i.e., interpersonal transgression, trust violation, private/public dimension), can be experienced at various degrees (i.e., trace to extreme), and can lead to multiple outcomes (i.e., make amends, get angry; Baumeister et al.,

1994; Nelissen & Zeelenberg, 2009). Those different experiences of guilt might vary in their behavioural expressions, occasionally presented as movements associated with pain or self-contempt (Keltner & Buswell, 1996) or gaze avoidance (Yu et al., 2017), but with specific core movements different from the previous conceptualisation. The studies presented in this thesis examined the production of facial signals of guilt in two different social contexts (Chapters 2 & 6) and resulted in the production of different facial signals associated with guilt. The difference in social partner, i.e., either someone that can be perceived as higher ranking (the experimenter in Chapter 2) or as a peer (the other player in Chapter 6), might have impacted the facial behaviours produced in association with feelings of guilt. Some secondary signals were produced by participants in both contexts but the main signals characterising the levels of reported self-guilt differed.

Finally, the judgement studies presented in this thesis present some linguistic limitations. Even if there are differences in the appraisal and behavioural outcomes between shame and guilt, it has been previously shown that English speakers use “guilt” and “shame” interchangeably (Fessler, 2004). Moreover, based on linguistic data different kinds of guilt were distinguished in Asian cultures (Bedford & Hwang, 2003), some that are not even defined in Western cultures, usually referring to the reason to feel guilty. I tried to overcome this issue in Chapter 4 by using different methodologies, but all experiments were conducted in English.

## Future directions

Although the traditional view states that facial displays should be understood on its own, in the absence of context (Ekman & Keltner, 1970), recent research has argued for a more ecological approach to judgement research (Matsumoto & Sung Hwang, 2010). Facial display is rarely occurring without context in real life (Amodio et al., 2007; Fernandez-Dols et al., 1993; Hess et al., 2016) and different cultures might associate different signals to similar contexts (Crivelli et al., 2016, 2017; Jack et al., 2016). To understand better how facial displays are processed, and more specifically facial signals associated with guilt, we could benefit from replicating the study presented in Chapter 4 and include the context each individual was in. If we want to gain insight into how faces are perceived in real life, it is important to design experiments that approach the complexity of real-life interactions. It would require more extensive samples, as more variables would be introduced, and the use of more complex analytical tool, allowing for such complexity in the design (e.g., temporal analysis using cross-recurrence quantification analysis - CRQA, Coco & Dale, 2014; lens modelling, Scherer et al., 2013).

Indeed, a more in-depth analysis would be an interesting step to further break down the mechanisms associated with guilt. The guilt stimuli collected in studies 1 and 6 (Chapters 1 & 6) present great complexity and diversity between individuals. A lens modelling approach might provide further information regarding the non-verbal signals associated with guilt, as well as identifying individual or cultural variations. Such model would help identify all the factors that could influence emotional perception, such as ethnicity and cultural but also presence of piercings, hair colour, hair length, face morphology, and clothes for instance. Because a face is never presented on its own in real life, it is important to acknowledge that everything about a person will affect the perception and interpretation of facial expressions. Regarding the perception of facial displays, it would also be interesting to take the analysis one step further. In



this thesis, I used simple frame-by-frame analysis of the expressions. A cross-recurrence quantification analysis might highlight more nuances in the expressions associated with guilt regarding the temporal succession of movements produced. Moreover, in Study 2 (Chapter 2), I collected extensive pinpoint data identifying genuine dynamic facial displays associated with various emotional states (discomfort, embarrassment, guilt, and surprise). It would be interesting to run a similar study to the one presented in Chapter 4, using all the stimuli associated with all the identified facial displays. Doing so would allow us to collect extensive data on the perception of dynamic, genuine, facial displays associated with secondary emotions often mistaken for guilt. We would thus be able to map out on the dimensional space the different emotional experience and gain insight into what makes the guilt displays stand out, or provide information as to why people tend to confuse guilt and embarrassment. The studies presented in this thesis remain exploratory and more work is needed to get a better understanding of the production and perception of a facial display of guilt.

As mentioned before, guilt can emerge in different situations, can be experienced at various degrees, and can lead to multiple outcomes (i.e., make amends, get angry; Baumeister et al., 1994; Nelissen & Zeelenberg, 2009). Its experience will also depend on people's propensity to experience guilt and on the rigidity of their moral compass (Cohen et al., 2012; Cohen et al., 2011). Different experiences of guilt might lead to different behavioural expressions (e.g., Keltner & Buswell, 1996; Yu et al., 2017). Investigating the facial signals of guilt in various contexts might help identify specific core movements reliably associated with self-reported feelings of guilt, as well as investigating the perceived signal of guilt in various situations. Moreover, exploring guilt in different situations might help understand how personality and felt guilt affect each other.

Looking at different contexts might be insightful regarding the behavioural outcomes associated with guilt. To understand further the cognitive mechanisms engaged with the experience of guilt, it would be interesting to investigate the production and perception of such signals taking a developmental approach. Investigating at what age children can express a facial signal of guilt would give information regarding the cognitive abilities required for this. It would be safe to assume such production would emerge after the development of secondary emotions (Lewis et al., 1989; Tangney, 1999), but we do not know how soon after the emergence of a feeling one can reliably produce associated facial signals. Also, does producing facial signals associated with guilt means being able to recognise said emotion on someone else's face? This area of research would bring valuable insight into the cognitive mechanisms needed to interpret accurately complex facial expressions.

Finally, expanding the sampled population and recruiting in various countries, differing in their cultural heritage, would be valuable to further study the impact of culture on both the production and perception of facial signals of guilt. Moreover, it would allow testing for the impact of religious beliefs and religious influences on the propensity to feel and express guilt.

## Conclusions

By studying facial signals associated with secondary emotions using a bottom-up approach, I contribute to the study of the communicative function of a facial signal of guilt. My studies tend to demonstrate that individuals produce a set of facial movements when feeling guilty, varying between contexts. Moreover, I have shown that observers from different countries seem to use similar facial signals to identify guilt on the face. Those signals elicit a functional response (punishing the wrongdoer) from the victims, in a way that is flexible depending on the relationship between the wrongdoer and the victim. I demonstrate the importance of contextual information when interpreting guilt signals as well as the impact of the methodology used on the ability to discriminate guilt signals from other emotional signals. With this set of experiments, I build on previous research on production and perception of facial signals and contribute to our understanding of the social value of guilt. Our findings could encourage researchers to try map facial displays using broader approaches and to favour a bottom-up methodology, putting the faces back in context.

## References

- Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2007). A dynamic model of guilt implications for motivation and self-regulation in the context of prejudice. *Psychological Science*, *18*(6), 524-530.
- Aviezer, H., Hassin, R., Bentin, S., & Trope, Y. (2008). Putting facial expressions back in context. *First impressions*, 255-286.
- Bandes, S. A. (2016). Remorse and criminal justice. *Emotion Review*, *8*(1), 14-19.
- Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on psychological science*, *1*(1), 28-58.
- Barrett, L. F. (2017a). *How emotions are made: The secret life of the brain*: Houghton Mifflin Harcourt.
- Barrett, L. F. (2017b). The theory of constructed emotion: an active inference account of interoception and categorization. *Social cognitive and affective neuroscience*, *12*(1), 1-23.
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1-68.
- Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science*, *20*(5), 286-290.
- Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annu. Rev. Psychol.*, *58*, 373-403.
- Barton, K., & Barton, M. K. (2019). Package 'MuMIn'. *R package version*, *1*(6).
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology letters*, *2*(3), 412-414.
- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: an interpersonal approach. *Psychological bulletin*, *115*(2), 243.
- Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., & Tapp, R. (2014). Featural processing in recognition of emotional facial expressions. *Cognition & Emotion*, *28*(3), 416-432.
- Bedford, O. A. (2004). The individual experience of guilt and shame in Chinese culture. *Culture & Psychology*, *10*(1), 29-52.
- Bedford, O. A., & Hwang, K. K. (2003). Guilt and Shame in Chinese Culture: A Cross-cultural Framework from the Perspective of Morality and Identity. *Journal for the Theory of Social Behaviour*, *33*(2), 127-144.
- Benedict, R. (1946). *Chrysanthemum and the Sword*. Patterns of Japanese Culture, Cleveland, New York (The World Publishing Company) 1946.
- Blair, R. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *358*(1431), 561-572.
- Blais, C., Fiset, D., Roy, C., Saumure Régimbald, C., & Gosselin, F. (2017). Eye fixation patterns for categorizing static and dynamic facial expressions. *Emotion*, *17*(7), 1107.
- Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *Plos One*, *3*(8), e3022.
- Borod, J. C., Haywood, C. S., & Koff, E. (1997). Neuropsychological aspects of facial asymmetry during emotional expression: A review of the normal adult literature. *Neuropsychology review*, *7*(1), 41-60.
- Boyd, R., & Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1533), 3281-3288.

- Bradbury, J. W., & Vehrencamp, S. L. (1998). Principles of animal communication.
- Brown, D. E. (1991). *Human universals*: McGraw-Hill New York.
- Campos, J., Barrett, K., Lamb, M., Goldsmith, H., & Stenberg, C. (1983). Socioemotional development In MM Haith & JJ Campos (Eds.), PH Mussen (Series Ed.), Handbook of child psychology: Vol. 2. Infancy and developmental psychobiology (pp. 783-915): New York: Wiley.
- Camras, L. A., Bakeman, R., Chen, Y., Norris, K., & Cain, T. R. (2006). Culture, ethnicity, and children's facial expressions: A study of European American, mainland Chinese, Chinese American, and adopted Chinese girls. *Emotion*, 6(1), 103-114. doi: 10.1037/1528-3542.6.1.103
- Carroll, J. (1985). *Guilt: the grey eminence behind character, history, and culture*: Routledge & Kegan Paul Books.
- Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of personality and social psychology*, 70(2), 205.
- Castelfranchi, C., & Poggi, I. (1990). Blushing as a discourse: Was Darwin wrong. *Shyness and embarrassment: Perspectives from social psychology*, 230-251.
- Chance, M. R. (1962). *An interpretation of some agonistic postures; the role of "cut-off" acts and postures*. Paper presented at the Symposia of the Zoological Society of London.
- Chen, C., & Jack, R. E. (2017). Discovering cultural differences (and similarities) in facial expressions of emotion. *Current opinion in psychology*, 17, 61-66.
- Christakis, N. A., & Fowler, J. H. (2010). *Connected: The surprising power of our social networks and how they shape our lives*: Little, Brown.
- Clark, C. (1990). Emotions and micropolitics in everyday life: Some patterns and paradoxes of "place." *Research agendas in the sociology of emotions*, 305-333.
- Coco, M. I., & Dale, R. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: an R package. *Frontiers in psychology*, 5, 510.
- Cohen, T. R., Panter, A. T., & Turan, N. (2012). Guilt Proneness and Moral Character. *Current Directions in Psychological Science*, 21(5), 355-359.
- Cohen, T. R., Wolf, S. T., Panter, A. T., & Insko, C. A. (2011). Introducing the GASP scale: a new measure of guilt and shame proneness. *Journal of personality and social psychology*, 100(5), 947.
- Cohn, J. F., & Schmidt, K. (2003). The timing of facial motion in posed and spontaneous smiles *Active Media Technology* (pp. 57-69): World Scientific.
- Cole, P. M., Zahn-Waxler, C., Fox, N. A., Usher, B. A., & Welsh, J. D. (1996). Individual differences in emotion regulation and behavior problems in preschool children. *Journal of Abnormal Psychology*, 105(4), 518.
- Cordaro, D. T., Sun, R., Keltner, D., Kamble, S., Huddar, N., & McNeil, G. (2018). Universals and cultural variations in 22 emotional expressions across five cultures. *Emotion*, 18(1), 75.
- Cousins, S. D. (1989). Culture and self-perception in Japan and the United States. *Journal of personality and social psychology*, 56(1), 124.
- Cowen, A. S., & Keltner, D. (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*, 114(38), E7900-E7909.
- Cowen, A. S., Sauter, D., Tracy, J. L., & Keltner, D. (2019). Mapping the Passions: Toward a High-Dimensional Taxonomy of Emotional Experience and Expression. *Psychological Science in the Public Interest*, 20(1), 69-90.
- Crivelli, C., & Fridlund, A. J. (2018). Facial displays are tools for social influence. *Trends in cognitive sciences*.
- Crivelli, C., & Fridlund, A. J. (2019). Inside-Out: From Basic Emotions Theory to the Behavioral Ecology View. *Journal of Nonverbal behavior*, 1-34.

- Crivelli, C., Russell, J. A., Jarillo, S., & Fernández-Dols, J.-M. (2016). The fear gasping face as a threat display in a Melanesian society. *Proceedings of the National Academy of Sciences*, *113*(44), 12403-12407.
- Crivelli, C., Russell, J. A., Jarillo, S., & Fernández-Dols, J.-M. (2017). Recognizing spontaneous facial expressions of emotion in a small-scale society of Papua New Guinea. *Emotion*, *17*(2), 337.
- Cryder, C. E., Springer, S., & Morewedge, C. K. (2012). Guilty feelings, targeted actions. *Personality and Social Psychology Bulletin*, *38*(5), 607-618.
- Daly, E. M., Lancee, W. J., & Polivy, J. (1983). A conical model for the taxonomy of emotional experience. *Journal of personality and social psychology*, *45*(2), 443.
- Darwall, S. L. (2006). *The second-person standpoint: Morality, respect, and accountability*: Harvard University Press.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: Albermarle.
- Davidson, R. J., Scherer, K. R., & Goldsmith, H. H. (2009). *Handbook of affective sciences*: Oxford University Press.
- Davila-Ross, M., Jesus, G., Osborne, J., & Bard, K. A. (2015). Chimpanzees (*Pan troglodytes*) produce the same types of 'laugh faces' when they emit laughter and when they are silent. *Plos One*, *10*(6), e0127337.
- De Hooge, I. E., Nelissen, R., Breugelmans, S. M., & Zeelenberg, M. (2011). What is moral about guilt? Acting "prosocially" at the disadvantage of others. *Journal of personality and social psychology*, *100*(3), 462.
- De Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. (2007). Moral sentiments and cooperation: Differential influences of shame and guilt. *Cognition and Emotion*, *21*(5), 1025-1042.
- de Jong, P. J., & Dijk, C. (2013). Social effects of facial blushing: influence of context and actor versus observer perspective. *Social and Personality Psychology Compass*, *7*(1), 13-26.
- De Jong, P. J., Peters, M. L., & De Cremer, D. (2003). Blushing may signify guilt: Revealing effects of blushing in ambiguous social situations. *Motivation and emotion*, *27*(3), 225-249.
- De Leersnyder, J. (2017). Emotional acculturation: a first review. *Current opinion in psychology*, *17*, 67-73.
- De Leersnyder, J., & Mesquita, B. (2015). How salient cultural concerns shape emotions: A behavioral coding study on biculturals' emotional frame switching.
- De Leersnyder, J., Mesquita, B., & Kim, H. S. (2011). Where do my emotions belong? A study of immigrants' emotional acculturation. *Personality and Social Psychology Bulletin*, *37*(4), 451-463.
- de Waal, F. (2011). What is an animal emotion? *Annals of the New York Academy of Sciences*, *1224*(1), 191-206.
- Delis, I., Chen, C., Jack, R. E., Garrod, O. G., Panzeri, S., & Schyns, P. G. (2016). Space-by-time manifold representation of dynamic facial expressions for emotion categorization. *Journal of vision*, *16*(8), 14-14.
- Dezecache, G., Mercier, H., & Scott-Phillips, T. C. (2013). An evolutionary approach to emotional communication. *Journal of Pragmatics*, *59*, 221-233.
- Dibble, J. L., Levine, T. R., & Park, H. S. (2012). The Unidimensional Relationship Closeness Scale (URCS): Reliability and validity evidence for a new measure of relationship closeness. *Psychological Assessment*, *24*(3), 565.
- Duchenne, B. (1862/1990). *The mechanism of human facial expression* Cambridge: Cambridge University Press.
- Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes: scanning emotional facial expressions. *Emotion*, *11*(4), 860.
- Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M., & Reno, R. R. (1989). Relation of sympathy and personal distress to prosocial behavior: a multimethod study. *Journal of personality and social psychology*, *57*(1), 55.

- Ekman, P. (1971). *Universals and cultural differences in facial expressions of emotion*. Paper presented at the Nebraska symposium on motivation.
- Ekman, P. (1980). Asymmetry in facial expression. *Science*, 209(4458), 833-834.
- Ekman, P. (1992a). Are there basic emotions?
- Ekman, P. (1992b). An argument for basic emotions. *Cognition & Emotion*, 6(3-4), 169-200.
- Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, 3(4), 364-370.
- Ekman, P., & Davidson, R. J. (1994). *The nature of emotion: Fundamental questions*: Oxford University Press.
- Ekman, P., & Friesen, W. V. (1969). The repertoire of non-verbal behavior: Categories, origins, usage and coding. *Semiotica*, 1(1), 49-98.
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), 124.
- Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of communication*, 22(4), 353-374.
- Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto: Consulting Psychologists Press.
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002). Facial action coding system—investigator's guide. *Research Nexus, Salt Lake City*.
- Ekman, P., & Keltner, D. (1970). Universal facial expressions of emotion. *California Mental Health Research Digest*, 8(4), 151-158.
- Ekman, P., & Rosenberg, E. L. (1997). *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*: Oxford University Press, USA.
- Elfenbein, H. A., & Ambady, N. (2003). When familiarity breeds accuracy: cultural exposure and facial emotion recognition. *Journal of personality and social psychology*, 85(2), 276.
- Elfenbein, H. A., Beaupré, M., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: cultural differences in the expression and recognition of posed facial expressions. *Emotion*, 7(1), 131.
- Emde, R. N., Johnson, W. F., & Easterbrooks, M. A. (1987). The do's and don'ts of early moral development: Psychoanalytic tradition and current research. *The emergence of morality in young children*, 245-276.
- Ferguson, T. J., & Crowley, S. L. (1997). Gender differences in the organization of guilt and shame. *Sex Roles*, 37(1-2), 19-44.
- Fernandez-Dols, J.-M. (2017). Natural Facial Expression. In J.-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 457-475). Oxford University Press.
- Fernandez-Dols, J. M., Sierra, B., & Ruiz-Belda, M. A. (1993). On the clarity of expressive and contextual information in the recognition of emotions: A methodological critique. *European Journal of Social Psychology*, 23(2), 195-202.
- Fessler, D. (1999). Toward an understanding of the universality of second order emotions. *Beyond nature or nurture: Biocultural approaches to the emotions*, 75-116.
- Fessler, D. (2004). Shame in two cultures: Implications for evolutionary approaches. *Journal of Cognition and Culture*, 4(2), 207-262.
- Field, T. M., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expression by neonates. *Science*, 218(4568), 179-181.
- Fontaine, J. R. J., & Scherer, K. R. (2013). Emotion is for doing: the action tendency component *Components of emotional meaning: A sourcebook* (pp. 170-185): Oxford University Press.
- Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, 18(12), 1050-1057.
- Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic behavior*, 6(3), 347-369.

- Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, 7(11), 1325-1330.
- Fridlund, A. J. (1994). *Human facial expression - An evolutionary view*. London: Academic Press.
- Fridlund, A. J. (2017). The behavioral ecology view of facial displays: 25 years later. In J.-M. F.-D. J. A. Russell (Ed.), *The science of facial expression* (pp. 77-92): Oxford University Press.
- Friesen, W. V., & Ekman, P. (1983). EMFACS-7: Emotional facial action coding system. *Unpublished manuscript, University of California at San Francisco*, 2(36), 1.
- Frijda, N. H. (1986). *The emotions: Studies in emotion and social interaction*. New-York: Cambridge University Press.
- Gächter, S., Herrmann, B., & Thöni, C. (2010). Culture and cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1553), 2651-2661.
- Galati, D., Sini, B., Schmidt, S., & Tinti, C. (2003). Spontaneous facial expressions in congenitally blind and sighted children aged 8-11. *Journal of Visual Impairment and Blindness*, 97(7), 418-428.
- Ganchrow, J. R., Steiner, J. E., & Daher, M. (1983). Neonatal facial expressions in response to different qualities and intensities of gustatory stimuli. *Infant Behavior and Development*, 6(2-3), 189-200.
- Gendron, M. (2017). Revisiting diversity: cultural variation reveals the constructed nature of emotion perception. *Current opinion in psychology*, 17, 145-150.
- Gendron, M., Crivelli, C., & Barrett, L. F. (2018). Universality reconsidered: Diversity in making meaning of facial expressions. *Current Directions in Psychological Science*, 27(4), 211-219.
- Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental science*, 14(5), 1012-1020.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), 504-528.
- Gross, J. J., & Barrett, L. F. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review*, 3(1), 8-16.
- Haidt, J. (2003). The moral emotions. In R. J. Davidson, K. R. Schere & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (Vol. 11, pp. 852-870). Oxford University Press.
- Haley, K. J., & Fessler, D. M. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, 26, 245-256.
- Hamilton, W. D. (1964). The genetical evolution of social behaviour. II. *Journal of theoretical Biology*, 7(1), 17-52.
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication methods and measures*, 1(1), 77-89.
- Henrich, J., & Boyd, R. (2001). Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of theoretical Biology*, 208(1), 79-89.
- Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., . . . Henrich, N. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, 327(5972), 1480-1484.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3), 61-83.
- 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(2), 280.
- Hess, U., Blaison, C., & Kafetsios, K. (2016). Judging facial emotion expressions in context: The influence of culture and self-construal orientation. *Journal of Nonverbal behavior*, 40(1), 55-64.
- Higgins, E. T. (1987). Self-discrepancy: a theory relating self and affect. *Psychological Review*, 94(3), 319.



- Hill, R. A., & Dunbar, R. I. (2003). Social network size in humans. *Human Nature, 14*(1), 53-72.
- Hinde, R. A. (1985). Was 'The Expression of the Emotions' a misleading phrase? *Animal Behaviour, 33*, 985-992.
- Hjortsjo, C. (1970). Man's Face and Mimic. *Language*.
- Horstmann, G. (2003). What do facial expressions convey: Feeling states, behavioral intentions, or actions requests? *Emotion, 3*(2), 150.
- Hothorn, T., Bretz, F., Westfall, P., Heiberger, R. M., Schuetzenmeister, A., Scheibe, S., & Hothorn, M. T. (2017). Package 'multcomp'. Obtenido de <http://cran.r-project.org/web/packages/multcomp/multcomp>.
- Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological monographs, 54*(2), 187-211.
- IBM, C. (2016). SPSS for Windows, version 24: IBM Corp Armonk (NY).
- Izard, C. E. (1977). *Human emotions*. New-York: Plenum Press.
- Izard, C. E. (1994). Innate and universal facial expressions: evidence from developmental and cross-cultural research.
- Izard, C. E. (1997). Emotions and facial expressions: A perspective from Differential Emotions Theory. *The psychology of facial expression, 2*, 57-77.
- Izard, C. E., Ackerman, B. P., & Schultz, D. (1999). Independent emotions and consciousness: Self-consciousness and dependent emotions.
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology, 19*(18), 1543-1548.
- Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General, 141*(1), 19.
- Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences, 201200155*.
- Jack, R. E., Sun, W., Delis, I., Garrod, O. G., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion. *Journal of Experimental Psychology: General, 145*(6), 708.
- Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1553), 2635-2650.
- Jensen, K., Hare, B., Call, J., & Tomasello, M. (2006). What's in it for me? Self-regard precludes altruism and spite in chimpanzees. *Proceedings of the Royal Society B: Biological Sciences, 273*(1589), 1013-1021.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy. *Handbook of personality: Theory and research, 3*(2), 114-158.
- Jonason, P. K., & Webster, G. D. (2010). The Dirty Dozen: A Concise Measure of the Dark Triad. *Psychological Assessment, 22*(2), 420-432.
- Jones, W. H., & Kugler, K. (1993). Interpersonal correlates of the Guilt Inventory. *Journal of personality assessment, 61*(2), 246-258.
- Julle-Danière, E., Whitehouse, J., Harris, C., Chung, M., Vrij, A., & Gustafsson, E. (in prep). Guilt outside of context.
- Julle-Danière, E., Whitehouse, J., Vrij, A., Gustafsson, E., & Waller, B. M. (in prep). The social outcomes of experiencing and seeing guilt.
- Julle-Danière, E., Whitehouse, J., Vrij, A., Gustafsson, E., & Waller, B. M. (under review). Are there non-verbal signals of guilt? *Plos One*.
- Keller, H., & Otto, H. (2009). The Cultural Socialization of Emotion Regulation During Infancy. *Journal of Cross-Cultural Psychology, 40*(6), 996-1011.
- Kelly, D. J., Liu, S., Rodger, H., Mielle, S., Ge, L., & Caldara, R. (2011). Developing cultural differences in face processing. *Developmental science, 14*(5), 1176-1184.
- Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of personality and social psychology, 68*(3), 441.

- Keltner, D., & Anderson, C. (2000). Saving face for Darwin: The functions and uses of embarrassment. *Current Directions in Psychological Science*, 9(6), 187-192.
- Keltner, D., & Buswell, B. N. (1996). Evidence for the distinctness of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. *Cognition & Emotion*, 10(2), 155-171.
- Keltner, D., & Gross, J. J. (1999). Functional accounts of emotions. *Cognition & Emotion*, 13(5), 467-480.
- Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional expression: Advances in basic emotion theory. *Journal of Nonverbal behavior*, 1-28.
- Keltner, D., Tracy, J., Sauter, D. A., Cordaro, D. C., & McNeil, G. (2016). Expression of emotion. *Handbook of emotions*, 467-482.
- Keltner, D., Tracy, J. L., Sauter, D., & Cowen, A. (2019). What Basic Emotion Theory Really Says for the Twenty-First Century Study of Emotion. *Journal of Nonverbal behavior*, 1-7.
- Ketelaar, T., & Au, W. T. (2003). The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction. *Cognition & Emotion*, 17(3), 429-453.
- Kim, H. S., Sherman, D. K., & Taylor, S. E. (2008). Culture and social support. *American Psychologist*, 63(6), 518.
- Krebs, J. R., & Davies, N. (1993). *An introduction to behavioural ecology*: Blackwell Scientific Publications.
- Krebs, J. R., & Dawkins, R. (1978). Animal signals: information or manipulation. In J. R. Krebs & R. Dawkins (Eds.), *Behavioural ecology: An evolutionary approach* (pp. 282-309): Blackwell Scientific Publications.
- Krippendorff, K. (1970). Bivariate agreement coefficients for reliability of data. *Sociological methodology*, 2, 139-150.
- Kroll, J., & Egan, E. (2004). Psychiatry, moral worry, and the moral emotions. *Journal of Psychiatric Practice*, 10(6), 352-360.
- Leary, M. R., Landel, J. L., & Patton, K. M. (1996). The motivated expression of embarrassment following a self-presentational predicament. *Journal of Personality*, 64(3), 619-636.
- LeDoux, J. (1996). *The Emotional Brain* (Vol. 384).
- Lee, K., Talwar, V., McCarthy, A., Ross, I., Evans, A., & Arruda, C. (2014). Can classic moral stories promote honesty in children? *Psychological Science*, 0956797614536401.
- Levenson, R. W. (2011). Basic emotion questions. *Emotion Review*, 3(4), 379-386.
- Lewis, M. (1993). Self-conscious emotions: Embarrassment, pride, shame, and guilt. In M. Lewis, JM Haviland (red.), *Handbook of emotions* (s. 563-573): Guilford Press, New York.
- Lewis, M. (1995a). Embarrassment: The emotion of self-exposure and evaluation.
- Lewis, M. (1995b). *Shame: The exposed self*: Simon and Schuster.
- Lewis, M., Alessandri, S. M., & Sullivan, M. W. (1992). Differences in shame and pride as a function of children's gender and task difficulty. *Child development*, 63(3), 630-638.
- Lewis, M., Sullivan, M. W., Stanger, C., & Weiss, M. (1989). Self development and self-conscious emotions. *Child development*, 146-156.
- Lien, J. J., Kanade, T., Cohn, J. F., & Li, C.-C. (1998). *Automated facial expression recognition based on FACS action units*. Paper presented at the Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition.
- Lien, J. J., Kanade, T., Cohn, J. F., & Li, C.-C. (2000). Detection, tracking, and classification of action units in facial expression. *Robotics and Autonomous Systems*, 31(3), 131-146.
- Longflier, L., Soussignan, R., Reissland, N., Leconte, M., Marret, S., Schaal, B., & Mellier, D. (2016). Emotional expressiveness of 5–6 month-old infants born very premature versus full-term at initial exposure to weaning foods. *Appetite*, 107, 494-500.
- Mangold, P. (1998). Interact [computer software]. *Arnstorf, Germany: Mangold International*.
- Marsh, A. A., Elenbaas, H. A., & Ambady, N. (2003). Nonverbal "accents" cultural differences in facial expressions of emotion. *Psychological Science*, 14(4), 373-376.

- Martin, P., & Bateson, P. (1993). *Measuring behaviour: an introductory guide*: Cambridge University Press.
- Martinez, L., Falvello, V. B., Aviezer, H., & Todorov, A. (2016). Contributions of facial expressions and body language to the rapid perception of dynamic emotions. *Cognition and Emotion, 30*(5), 939-952.
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior research methods, 44*(2), 314-324.
- Matsumoto, D. (1989). Cultural influences on the perception of emotion. *Journal of Cross-Cultural Psychology, 20*(1), 92-105.
- Matsumoto, D. (1992). American-Japanese cultural differences in the recognition of universal facial expressions. *Journal of Cross-Cultural Psychology, 23*(1), 72-84.
- Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M., & Frank, M. (2008). Facial expressions of emotion. *Handbook of emotions, 3*, 211-234.
- Matsumoto, D., & Sung Hwang, H. (2010). Judging faces in context. *Social and Personality Psychology Compass, 4*(6), 393-402.
- Matsumoto, D., Yoo, S. H., & Fontaine, J. (2008). Mapping expressive differences around the world the relationship between emotional display rules and individualism versus collectivism. *Journal of Cross-Cultural Psychology, 39*(1), 55-74.
- Matsumoto, D., Yoo, S. H., Hirayama, S., & Petrova, G. (2005). Development and validation of a measure of display rule knowledge: the display rule assessment inventory. *Emotion, 5*(1), 23.
- Mayr, E. (1961). Cause and effect in biology. *Science, 134*(3489), 1501-1506.
- McAlister, R., Harkness, E., & Nicoll, J. (1998). An ultrasound investigation of the lip levator musculature. *The European Journal of Orthodontics, 20*(6), 713-720.
- McDuff, D., Mahmoud, A., Mavadati, M., Amr, M., Turcot, J., & Kaliouby, R. e. (2016). *AFFDEX SDK: a cross-platform real-time multi-face expression recognition toolkit*. Paper presented at the Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems.
- Mehu, M., & Scherer, K. R. (2015). Emotion categories and dimensions in the facial communication of affect: An integrated approach. *Emotion, 15*(6), 798.
- Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions: a review. *Psychological bulletin, 112*(2), 179.
- Mesquita, B., & Leu, J. (2007). The cultural psychology of emotions. In S. Kitayama & D. Cohen (Eds.), *Handbook for cultural psychology* (pp. pp. 734-759). New York: NY: Guilford.
- Miceli, M., & Castelfranchi, C. (2018). Reconsidering the Differences Between Shame and Guilt. *Europe's Journal of Psychology, 14*(3), 710-733.
- Mielliet, S., Vizioli, L., He, L., Zhou, X., & Caldara, R. (2013). Mapping face recognition information use across cultures. *Frontiers in psychology, 4*, 34.
- Miossec, S., & Kheddar, A. (2009). *Human motion in cooperative tasks: Moving object case study*. Paper presented at the Robotics and Biomimetics, 2008. ROBIO 2008. IEEE International Conference on.
- Mohiyeddini, C., Bauer, S., & Semple, S. (2013). Displacement behaviour is associated with reduced stress levels among men but not women. *Plos One, 8*(2), e56355.
- Mohiyeddini, C., & Semple, S. (2013). Displacement behaviour regulates the experience of stress in men. *Stress, 16*(2), 163-171.
- Mohr, C., Jonauskaitė, D., Dan-Glauser, E. S., Uusküla, M., & Dael, N. (2018). Unifying research on colour and emotion: Time for a cross-cultural survey on emotion associations to colour terms. *Progress in Colour Studies: Cognition, language and beyond, 209-222*.
- Muris, P., & Meesters, C. (2014). Small or big in the eyes of the other: on the developmental psychopathology of self-conscious emotions as shame, guilt, and pride. *Clinical child and family psychology review, 17*(1), 19-40.
- Nelissen, R., & Zeelenberg, M. (2009). When guilt evokes self-punishment: evidence for the existence of a Dobby Effect. *Emotion, 9*(1), 118.

- Nelson, N. L., & Russell, J. A. (2013). Universality revisited. *Emotion Review*, 5(1), 8-15.
- O'Malley, M. N., & Greenberg, J. (1983). Sex differences in restoring justice: The down payment effect. *Journal of Research in Personality*, 17(2), 174-185.
- Ohbuchi, K.-i., Kameda, M., & Agarie, N. (1989). Apology as aggression control: its role in mediating appraisal of and response to harm. *Journal of personality and social psychology*, 56(2), 219.
- Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition*, 108(1), 222-231.
- Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, 97(3), 315.
- Oster, H. (2006). Baby FACS: Facial Action Coding System for infants and young children. *Unpublished monograph and coding manual*. New York University.
- Palan, S., & Schitter, C. (2018). Prolific. ac—a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22-27.
- Panksepp, J. (2007). Criteria for basic emotions: Is DISGUST a primary “emotion”? *Cognition and Emotion*, 21(8), 1819-1828.
- Parkinson, B. (2005). Do facial movements express emotions or communicate motives? *Personality and Social Psychology Review*, 9(4), 278-311.
- Parkinson, B., & Illingworth, S. (2009). Guilt in response to blame from others. *Cognition and Emotion*, 23(8), 1589-1614.
- Parr, L., & Waller, B. (2006). The evolution of human emotion *Evolution of nervous systems: a comprehensive reference*: Academic Press Inc.
- Piaget, J. (1965). 1932. The Moral Judgment of the Child: Free Press, New York.
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion *Theories of emotion* (pp. 3-33): Elsevier.
- Qualtrics, s. (2012). Qualtrics.
- Rakoczy, H. (2008). Taking fiction seriously: Young children understand the normative structure of joint pretence games. *Developmental Psychology*, 44(4), 1195.
- Rebega, O. L., Apostol, L., Benga, O., & Miclea, M. (2013). Inducing Guilt: A Literature Review. *Procedia-Social and Behavioral Sciences*, 78, 536-540.
- Reed, L. I., Zeglen, K. N., & Schmidt, K. L. (2012). Facial expressions as honest signals of cooperative intent in a one-shot anonymous Prisoner's Dilemma game. *Evolution and Human Behavior*, 33(3), 200-209.
- Rodger, H., Kelly, D. J., Blais, C., & Caldara, R. (2010). Inverting faces does not abolish cultural diversity in eye movements. *Perception*, 39(11), 1491-1503.
- Rosenstock, S., & O'Connor, C. (2018). When It's Good to Feel Bad: An Evolutionary Model of Guilt and Apology. *Frontiers in Robotics and AI*, 5(9). doi: 10.3389/frobt.2018.00009
- Roy, C., Blais, C., Fiset, D., Rainville, P., & Gosselin, F. (2015). Efficient information for recognizing pain in facial expressions. *European Journal of Pain*, 19(6), 852-860.
- Russell, J. A. (1989). Measures of emotion. In R. P. H. Kellerman (Ed.), *Emotion: Theory, research, and experience* (Vol. 4, pp. 83-112): San Diego, CA: Academic.
- Russell, J. A. (1991). Culture and the categorization of emotions. *Psychological bulletin*, 110(3), 426.
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological bulletin*, 115(1), 102.
- Russell, J. A., & Fernández-Dols, J. M. (1997). What does a facial expression mean. *The psychology of facial expression*, 1.
- Samson, A. V., & Waller, B. M. (2010). Not growling but smiling: New interpretations of the bared-teeth motif in the pre-Columbian Caribbean. *Current Anthropology*, 51(3), 425-433.
- Scherer, K. R., Mortillaro, M., & Mehu, M. (2013). Understanding the mechanisms underlying the production of facial expression of emotion: A componential perspective. *Emotion Review*, 5(1), 47-53.

- Schmidt, K. L., Bhattacharya, S., & Denlinger, R. (2009). Comparison of deliberate and spontaneous facial movement in smiles and eyebrow raises. *Journal of Nonverbal behavior*, *33*(1), 35-45.
- Schmidt, K. L., & Cohn, J. F. (2001). Human facial expressions as adaptations: Evolutionary questions in facial expression research. *American journal of physical anthropology*, *116*(S33), 3-24.
- Schmidt, K. L., Cohn, J. F., & Tian, Y. (2003). Signal characteristics of spontaneous facial expressions: Automatic movement in solitary and social smiles. *Biological psychology*, *65*(1), 49-66.
- Schmidt, M. F., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *Plos One*, *6*(10), e23223.
- Schug, J., Matsumoto, D., Horita, Y., Yamagishi, T., & Bonnet, K. (2010). Emotional expressivity as a signal of cooperation. *Evolution and Human Behavior*, *31*(2), 87-94.
- Schurgin, M., Nelson, J., Iida, S., Ohira, H., Chiao, J., & Franconeri, S. (2014). Eye movements during emotion recognition in faces. *Journal of vision*, *14*(13), 14-14.
- Schyns, P. G. (1998). Diagnostic recognition: task constraints, object information, and their interactions. *Cognition*, *67*(1-2), 147-179.
- Seyfarth, R. M., & Cheney, D. L. (2012). The evolutionary origins of friendship. *Annual review of psychology*, *63*, 153-177.
- Sgoifo, A., Braglia, F., Costoli, T., Musso, E., Meerlo, P., Ceresini, G., & Troisi, A. (2003). Cardiac autonomic reactivity and salivary cortisol in men and women exposed to social stressors: relationship with individual ethological profile. *Neuroscience & Biobehavioral Reviews*, *27*(1-2), 179-188.
- Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for? *Current Directions in Psychological Science*, *20*(6), 395-399.
- Shuman, V., Clark-Polner, E., Meuleman, B., Sander, D., & Scherer, K. R. (2017). Emotion perception from a componential perspective. *Cognition and Emotion*, *31*(1), 47-56.
- Siegel, E. H., Sands, M. K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., . . . Barrett, L. F. (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological bulletin*, *144*(4), 343.
- Silk, J. B. (2005). Practicing Hamilton's rule: kin selection in primate groups *Cooperation in primates and humans* (pp. 25-46): Springer.
- Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, *16*(3), 184-189.
- Smith, M. L., Grün, D., Bevitt, A., Ellis, M., Ciripan, O., Scrimgeour, S., . . . Ewing, L. (2018). Transmitting and decoding facial expressions of emotion during healthy aging: More similarities than differences. *Journal of vision*, *18*(9), 10-10.
- Snijders, T. A., & Borgatti, S. P. (1999). Non-parametric standard errors and tests for network statistics. *Connections*, *22*(2), 161-170.
- Soussignan, R., Dollion, N., Schaal, B., Durand, K., Reissland, N., & Baudouin, J.-Y. (2018). Mimicking emotions: how 3–12-month-old infants use the facial expressions and eyes of a model. *Cognition and Emotion*, *32*(4), 827-842.
- Sperber, D., & Baumard, N. (2012). Moral reputation: An evolutionary and cognitive perspective. *Mind & Language*, *27*(5), 495-518.
- Steiner, J. E. (1979). Human facial expressions in response to taste and smell stimulation. *Advances in child development and behavior*, *13*, 257-295.
- Stratou, G., Van Der Schalk, J., Hoegen, R., & Gratch, J. (2017). *Refactoring facial expressions: An automatic analysis of natural occurring facial expressions in iterative social dilemma*. Paper presented at the 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII).
- Tajfel, H. E. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*: Academic Press.

- Tajfel, H. E., Turner, J. C., Austin, W. G., & Worchel, S. (1979). An integrative theory of intergroup conflict. *Organizational identity: A reader*, 56-65.
- Talwar, V., Yachison, S., & Leduc, K. (2015). Promoting Honesty: The Influence of Stories on Children's Lie-Telling Behaviours and Moral Understanding. *Infant and Child Development*.
- Tangney, J. P. (1991). Moral affect: The good, the bad, and the ugly. *Journal of Personality and Social Psychology* ( 61), 598-607.
- Tangney, J. P. (1999). The self-conscious emotions: Shame, guilt, embarrassment and pride.
- Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annu. Rev. Psychol.*, 58, 345-372.
- Tangney, J. P., & Tracy, J. L. (2012). Self-conscious emotions. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 446-478). New York: Guilford Press.
- Taylor, G. (1996). Guilt and remorse. *The emotions: Social, cultural and biological dimensions*, 57-73.
- Team, R. C. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013: ISBN 3-900051-07-0.
- Teroni, F., & Deonna, J. A. (2008). Differentiating shame from guilt. *Consciousness and cognition*, 17(3), 725-740.
- Thamm, R. A. (2006). The classification of emotions *Handbook of the sociology of emotions* (pp. 11-37): Springer.
- Thornhill, R., & Gangestad, S. W. (1999). Facial attractiveness. *Trends in cognitive sciences*, 3(12), 452-460.
- Tignor, S. M., & Colvin, C. R. (2019). The meaning of guilt: Reconciling the past to inform the future. *Journal of personality and social psychology*, 116(6), 989.
- Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für tierpsychologie*, 20(4), 410-433.
- Tomasello, M. (2008). *Origins of human communication*: Cambridge, MA: MIT Press.
- Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual review of psychology*, 64, 231-255.
- Tracy, J. L., & Randles, D. (2011). Four models of basic emotions: a review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emotion Review*, 3(4), 397-405.
- Tracy, J. L., & Robins, R. W. (2004). " Putting the Self Into Self-Conscious Emotions: A Theoretical Model". *Psychological Inquiry*, 15(2), 103-125.
- Tracy, J. L., & Robins, R. W. (2008). The nonverbal expression of pride: evidence for cross-cultural recognition. *Journal of personality and social psychology*, 94(3), 516.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly review of biology*, 46(1), 35-57.
- Troisi, A. (2002). Displacement activities as a behavioral measure of stress in nonhuman primates and human subjects. *Stress*, 5(1), 47-54.
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*: Cambridge University Press.
- Turner, J. (2000). *On the origins of human emotions: A sociological inquiry into the evolution of human affect*: Stanford University Press.
- Vaidya, A. R., Jin, C., & Fellows, L. K. (2014). Eye spy: The predictive value of fixation patterns in detecting subtle and extreme emotions from faces. *Cognition*, 133(2), 443-456.
- Vaish, A. (2018). The prosocial functions of early social emotions: the case of guilt. *Current opinion in psychology*, 20, 25-29.
- Wagner, H. L., & Smith, J. (1991). Facial expression in the presence of friends and strangers. *Journal of Nonverbal behavior*, 15(4), 201-214.

- Waller, B. M., Bard, K. A., Vick, S.-J., & Smith Pasqualini, M. C. (2007). Perceived differences between chimpanzee (*Pan troglodytes*) and human (*Homo sapiens*) facial expressions are related to emotional interpretation. *Journal of Comparative Psychology, 121*(4), 398.
- Waller, B. M., & Micheletta, J. (2013). Facial expression in nonhuman animals. *Emotion Review, 5*(1), 54-59.
- Waller, B. M., Parr, L., Gothard, K. M., Burrows, A., & Fuglevand, A. J. (2008). Mapping the contribution of single muscles to facial movements in the rhesus macaque. *Physiology & Behavior, 95*(1-2), 93-100.
- Waller, B. M., Warmelink, L., Liebal, K., Micheletta, J., & Slocombe, K. E. (2013). Pseudoreplication: a widespread problem in primate communication research. *Animal Behaviour, 86*(2), 483-488.
- Waller, B. M., Whitehouse, J., & Micheletta, J. (2016). Macaques can predict social outcomes from facial expressions. *Animal Cognition, 19*(5), 1031-1036.
- Warnes, G. R., Bolker, B., Lumley, T., Warnes, M. G. R., & Imports, M. (2018). Package 'gmodels'.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology, 54*(6), 1063.
- Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological bulletin, 98*(2), 219.
- Weisman, M. R. (2014). *Showing remorse: Law and the social control of emotion*: Ashgate Publishing, Ltd.
- Whitehead, H. (2008). *Analyzing animal societies: quantitative methods for vertebrate social analysis*: University of Chicago Press.
- Wierzbicka, A. (1992). Defining emotion concepts. *Cognitive science, 16*(4), 539-581.
- Wyman, E., Rakoczy, H., & Tomasello, M. (2009). Normativity and context in young children's pretend play. *Cognitive development, 24*(2), 146-155.
- Yan, X., Young, A. W., & Andrews, T. J. (2017). Differences in holistic processing do not explain cultural differences in the recognition of facial expression. *The Quarterly Journal of Experimental Psychology, 70*(12), 2445-2459.
- Yik, M. S. (1999). Interpretation of faces: A cross-cultural study of a prediction from Fridlund's theory. *Cognition & Emotion, 13*(1), 93-104.
- Yu, H., Duan, Y., & Zhou, X. (2017). Guilt in the eyes: Eye movement and physiological evidence for guilt-induced social avoidance. *Journal of Experimental Social Psychology, 71*, 128-137.
- Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology, 43*(2), 303-311.
- Zemack-Rugar, Y., Bettman, J. R., & Fitzsimons, G. J. (2007). The effects of nonconsciously priming emotion concepts on behavior. *Journal of personality and social psychology, 93*(6), 927.

## Appendices

### Appendix 1. Personality and Demographic Questionnaires

#### Big Five – “How am I in general?”

People’s personality is commonly measured using five main traits or domains used to estimate individual differences in thinking, feeling, and: Extraversion, Agreeableness, Conscientiousness, Negative Emotionality (alternatively labelled Neuroticism vs. Emotional Stability), and Open Mindedness (alternatively labelled Openness to Experience). We chose to measure those personality traits using a questionnaire made of 44 items asking people to reflect on different aspect of their behaviours or personalities in their daily life– “How am I in general?”. On a scale from 1 (*Disagree strongly*) to 5 (*Agree strongly*), participants had to indicate how the presented characteristics applied to them.

**Scoring.** Reverse score the items labelled “R” and compute the scale score as the mean of the following items:

- Extraversion: 1, 6R, 11, 16, 21R, 26, 31R, 36
- Agreeableness: 2R, 7, 12R, 17, 22, 27R, 32, 37R, 42
- Conscientiousness: 3, 8R, 12, 19R, 23R, 28, 33, 38, 43R
- Neuroticism: 4, 9R, 14, 19, 24R, 29, 34R, 39
- Openness: 5, 10, 15, 20, 25, 30, 35R, 40, 41R, 44

*Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement.*

*I am someone who...*



1. *Is talkative*
2. *Tends to find fault with others*
3. *Does a thorough job*
4. *Is depressed, blue*
5. *Is original*
6. *Is reserved,*
7. *Is helpful and unselfish with others*
8. *Can be somewhat careless*
9. *Is relaxed, handles stress well*
10. *Is curious about many different things*
11. *Is full of energy*
12. *Starts quarrels with others*
13. *Is a reliable worker*
14. *Can be tense*
15. *Is ingenious, a deep thinker*
16. *Generated a lot of enthusiasm*
17. *Has a forgiving nature*
18. *Tends to be disorganised*
19. *Worries a lot*
20. *Has an active imagination*
21. *Tends to be quiet*
22. *Is generally trusting*
23. *Tends to be lazy*
24. *Is emotionally stable, not easily upset*
25. *Is inventive*
26. *Has an assertive personality*
27. *Can be cold and aloof*
28. *Perseveres until the task is finished*
29. *Can be moody*
30. *Values artistic, aesthetic experiences*
31. *Is sometimes shy, inhibited*
32. *Is considerate and kind to almost everyone*
33. *Does things efficiently*
34. *Remains calm in tense situations*
35. *Prefers work that is routine*
36. *Is outgoing, sociable*
37. *Is sometimes rude to others*
38. *Makes plans and follows through with them*
39. *Gets nervous easily*
40. *Likes to reflect, play with ideas*
41. *Has few artistic interests*
42. *Likes to cooperate with others*
43. *Is easily distracted*
44. *Is sophisticated in art, music, or literature*

## Dirty Dozen

The Dirty Dozen is a 12-item self-report scale measuring the Dark Triad: narcissism (e.g., the need for attention, associated with high self-perception), psychopathy (e.g., lack of remorse, callous behaviour tendencies), and Machiavellianism (e.g., manipulating people, unethical behaviours to reach one goal). The Dark Triad refer to three undesirable personality trait, often associated with social shortcoming and unethical decision-making. Using a scale ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*), participants had to indicate how much they agreed with each of the 12 statements. We used it to see how those personality traits could influence one self-reported guilt or judgement of guilt, as well as behavioural outcomes.

**Scoring.** Scores are computed by averaging the 4 items in each subscale.

- Machiavellianism: 1-4
- Narcissism: 5-8
- Psychopathy: 9-12

1. *I tend to manipulate others to get my way*
2. *I have used deceit or lied to get my way*
3. *I have use flattery to get my way*
4. *I tend to exploit others towards my own end.*
5. *I tend to lack remorse.*
6. *I tend to be unconcerned with the morality of my actions.*
7. *I tend to be callous or insensitive*
8. *I tend to be cynical.*
9. *I tend to want others to admire me.*
10. *I tend to want others to pay attention to me.*
11. *I tend to seek prestige or status*
12. *I tend to expect special favours from others.*

## PANAS

The Positive and Negative Affect Schedule scales are composed of two 10-item mood scales. Each scale is made of 10 emotion words that are most commonly used when describing an emotional state. The positive scale is composed of *active, alert, attentive, determined, enthusiastic, excited, inspired, interested, proud, and strong*; the negative scale is composed of *afraid, ashamed, distressed, guilty, hostile, irritable, jittery, nervous, scared, and upset*. Participants report their emotional state using a 5-point scale: 1.very slightly/not at all; 2. a little; 3.moderately; 4.quite a bit; 5. Extremely. The Positive Affect Score is calculated by adding the scores (1 to 5) of all the positive items; the Negative Affect Score is calculated by adding the scores (1 to 5) of all the negative items. When used with short-term instructions (e.g., *right now*), the PANAS scales measure fluctuation in mood and emotional states. In this study, we also looked at the evolution across time of the Positive and Negative Affect Scores as well as the changes in specific emotional state (in bold previously: proud, ashamed, distressed, and guilty). We used the PANAS to control for the efficiency of our experimental settings:

- Chapter 2: does the participant in Study 1 feel worst and report an increased level of guilt after the interaction with the researcher ?
- Chapter 6: does Player 1 feel generally worst and report an increase of guilty feeling after receiving the feedback of the game compare to before? Does receiving positive feedback regarding one's performance increase the positive mood of Player 2 and reduced the level of guilt compared to right after the game?

## GASP

The Guilt and Shame Proneness scale measures individual differences in the propensity to experience guilt and shame. The GASP contains two guilt subscales: negative behaviour-evaluations (i.e., feeling bad about how you acted, useful to detect potential unethical decision making) and repair action tendencies (i.e., behaviours and behavioural tendencies); and two shame subscales: negative self-evaluations (i.e., feeling bad about yourself, measuring the moral, prosocial aspect of shame proneness) and withdrawal action tendencies (i.e., action tendencies and behaviours).

Participants are presented with 16 items and asked, using a scale ranging from 1 (*Very Unlikely*) to 7 (*Very Likely*), how well each situation applied to them. We used the GASP in order to assess the relationship between guilt proneness and self-reported guilt as well as judged guilt. We wanted to investigate how one's propensity to experience guilt translated in the present setting.

**GASP scoring.** The GASP is scored by averaging the four items in each subscale.

- Guilt-Negative-Behavior-Evaluation (NBE): 1, 9, 14, 16
- Guilt-Repair: 2, 5, 11, 15
- Shame-Negative-Self-Evaluation (NSE): 3, 6, 10, 13
- Shame-Withdraw: 4, 7, 8, 12

*In this questionnaire you will read about situations that people are likely to encounter in day-to-day life, followed by common reactions to those situations. As you read each scenario, try to imagine yourself in that situation. Then indicate the likelihood that you would react in the way described.*

\_\_\_\_\_ 1. *After realizing you have received too much change at a store, you decide to keep it because the salesclerk doesn't notice. What is the likelihood that you would feel uncomfortable about keeping the money?*

\_\_\_\_\_ 2. *You are privately informed that you are the only one in your group that did not make the honor society because you skipped too many days of school. What is the likelihood that this would lead you to become more responsible about attending school?*

\_\_\_\_\_ 3. You rip an article out of a journal in the library and take it with you. Your teacher discovers what you did and tells the librarian and your entire class. What is the likelihood that this would make you would feel like a bad person?

\_\_\_\_\_ 4. After making a big mistake on an important project at work in which people were depending on you, your boss criticizes you in front of your coworkers. What is the likelihood that you would feign sickness and leave work?

\_\_\_\_\_ 5. You reveal a friend's secret, though your friend never finds out. What is the likelihood that your failure to keep the secret would lead you to exert extra effort to keep secrets in the future?

\_\_\_\_\_ 6. You give a bad presentation at work. Afterwards your boss tells your coworkers it was your fault that your company lost the contract. What is the likelihood that you would feel incompetent?

\_\_\_\_\_ 7. A friend tells you that you boast a great deal. What is the likelihood that you would stop spending time with that friend?

\_\_\_\_\_ 8. Your home is very messy and unexpected guests knock on your door and invite themselves in. What is the likelihood that you would avoid the guests until they leave?

\_\_\_\_\_ 9. You secretly commit a felony. What is the likelihood that you would feel remorse about breaking the law?

10. You successfully exaggerate your damages in a lawsuit. Months later, your lies are discovered and you are charged with perjury. What is the likelihood that you would think you are a despicable human being?

\_\_\_\_\_ 11. You strongly defend a point of view in a discussion, and though nobody was aware of it, you realize that you were wrong. What is the likelihood that this would make you think more carefully before you speak?

\_\_\_\_\_ 12. You take office supplies home for personal use and are caught by your boss. What is the likelihood that this would lead you to quit your job?

\_\_\_\_\_ 13. You make a mistake at work and find out a coworker is blamed for the error. Later, your coworker confronts you about your mistake. What is the likelihood that you would feel like a coward?

\_\_\_\_\_ 14. At a coworker's housewarming party, you spill red wine on their new creamcolored carpet. You cover the stain with a chair so that nobody notices your mess. What is the likelihood that you would feel that the way you acted was pathetic?

\_\_\_\_\_ 15. While discussing a heated subject with friends, you suddenly realize you are shouting though nobody seems to notice. What is the likelihood that you would try to act more considerately toward your friends?

\_\_\_\_\_ 16. You lie to people but they never find out about it. What is the likelihood that you would feel terrible about the lies you told?

## URCS

The Unidimensional Relationship Closeness Scale is a 12-item self-report scale measuring closeness of social and personal relationships. It is used to measure the strength of the emotional bond between two individuals, and applied to both romantic and friendly relationships. Using a scale ranging from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*), participants had to indicate how each statement applied to them. The URCS conceptualise closeness as a continuum, which allowed us to include this measure as a variable in our models. We used this score to investigate how closeness affected one's behavioural response, both as a guilty part or as a victim of one's wrongdoing.

**Scoring.** The items are averaged to create a single overall closeness score.

*The following questions refer to your relationship with **the other player**. Please think about your relationship with **the other player** when responding to the following questions.*

1. My relationship with **the other player** is close.
2. When we are apart, I miss **the other player** a great deal.
3. **The other player** and I disclose important personal things to each other.
4. **The other player** and I have a strong connection.
5. **The other player** and I want to spend time together.
6. I'm sure of my relationship with **the other player**.
7. **The other player** is a priority in my life.
8. **The other player** and I do a lot of things together.
9. When I have free time, I choose to spend it alone with **the other player**.
10. I think about **the other player** a lot.
11. My relationship with **the other player** is important in my life.
12. I consider **the other player** when making important decisions

## Acculturation Questionnaires

### Asian questionnaire (De Leersnyder et al., 2011)

A four-item scale measured the degree of social contact with members of the host culture; participants used a scale ranging from 1 to 5 to indicate how much they agree with each statement.

English language use: How would you describe your English language use?

scale ranging from 1 (*I speak my native language all the time*) to 5 (*I speak English all the time*)

Social network: How would you describe your social relationships' ethnicity? (friends, colleagues, and neighbours)

scale from 1 (*heritage culture only*) to 5 (*British/European only*)

### Western questionnaire (Delis et al., 2016)

Each potential observer completed the following questionnaire:

1. Have you ever lived in non-Western\* country before (e.g. on a gap year, summer work, move due parental employment)?
2. How many weeks have you spent in a non-Western country (e.g. on vacation)?
3. Have you ever dated or had a very close friendship with a non-Western person?
4. Have you ever been involved with any non-Western culture societies/groups?

\*by Western groups/countries, we are referring to Europe (East and West), USA, Canada, United Kingdom, Australia and New Zealand.

Appendix 2. SFEC 2015-097 Favourable opinion letter – Study 1 (Chapter 2)





**Science Faculty Ethics Committee**  
Science Faculty Office  
University of Portsmouth  
St Michael's Building  
White Swan Road  
PORTSMOUTH  
PO1 2DT

Eglantine Julle-Daniere  
*Department of Psychology*  
University of Portsmouth

T: 023 9284 3379  
[ethics-sci@port.ac.uk](mailto:ethics-sci@port.ac.uk)

11<sup>th</sup> January 2016

[Eglantine.julle-daniere@port.ac.uk](mailto:Eglantine.julle-daniere@port.ac.uk)

### **FAVOURABLE ETHICAL OPINION**

Study Title: **The behavioural expression of guilt**

Reference Number: **SFEC 2015-097** (Please quote this in any correspondence)

Thank you for re-submitting your application to the Science Faculty Ethics Committee (SFEC) dated 9<sup>th</sup> December 2015, in accordance with current procedures<sup>1</sup>

I am pleased to inform you that SFEC was content to grant a favourable ethical opinion of the above research on the basis described in the submitted documents listed at Annex A, and subject to standard general conditions<sup>2</sup>.

Please note that the favourable opinion of SFEC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including the University of Portsmouth or supervisor, prior to the start of the study.

Wishing you every success in your research

A handwritten signature in black ink, appearing to read 'S. Kolstoe'.

Dr Simon Kolstoe  
Alternate Vice Chair, Science Faculty Ethics Committee

#### **Information:**

*Dr Bridget Waller*

*Prof Aldert Vrije*

Holly Shawyer - Faculty Administrator

---

<sup>1</sup> Procedures for Ethical Review, Science Faculty Ethics Committee, University of Portsmouth, October 2012 (to be updated).

<sup>2</sup> After ethical review – Guidance for researchers (Please read).

Appendix 3. SFEC 2016-061 Favourable with Conditions– Study 2 (Chapters 2 & 3)



**Science Faculty Ethics Committee**  
Science Faculty Office  
University of Portsmouth  
St Michael's Building  
White Swan Road  
PORTSMOUTH  
PO1 2DT

Ms Eglantine Julle-Daniere  
Department of Psychology  
University of Portsmouth

T: 023 9284 3379  
[ethics-sci@port.ac.uk](mailto:ethics-sci@port.ac.uk)

[eglantine.julle-daniere@port.ac.uk](mailto:eglantine.julle-daniere@port.ac.uk)

2 August 2016

## **FAVOURABLE ETHICAL OPINION – WITH CONDITIONS**

**Study Title:** The behavioural expression of guilt - Can guilt be perceived by naïve observers?

**Reference Number:** SFEC 2016-061

**Date Resubmitted:** 25 July 2016

Thank you for submitting your protocol to the Science Faculty Ethics Committee (SEFC) for ethical review in accordance with current procedures<sup>1</sup>.

I am pleased to inform you that SFEC was content to grant a favourable ethical opinion of the above research on the basis described in the submitted documents listed at Annex A, and subject to standard general conditions (*See Annex B*), and the following specific minor conditions.

### Conditions/Comments

A - 1.3 - Start date 2016 (not 2017).

B - 10.4 - Shouldn't particular pool volunteers be offered the course credit OR £10, the latter especially if they have already gained enough course credits?

C - Poster - should state 'course credits or £10) in line with previous comments.

If you would find it helpful to discuss any of the matters raised above or seek further clarification from a member of the Committee, you are welcome to contact [ethics-sci@port.ac.uk](mailto:ethics-sci@port.ac.uk) who will circulate your queries to SFEC

Please note that the favourable opinion of SFEC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including the University of Portsmouth or supervisor, prior to the start of the study.

Wishing you every success in your research

---

<sup>1</sup> Procedures for Ethical Review, Science Faculty Ethics Committee, University of Portsmouth, October 2012 (to be updated).



Dr Paul Morris  
Chair, Science Faculty Ethics Committee/Vice Chair, Science Faculty Ethics Committee

### Annexes

A - Documents reviewed  
B - After ethical review - Guidance for researchers

### Information:

Dr Bridget Waller – PhD Supervisor  
Holly Shawyer - Faculty Administrator

### **Statement of compliance**

SFEC is constituted in accordance with the Governance Arrangements set out by the University of Portsmouth

### **After Ethical Review**

If unfamiliar, please consult the advice After Ethical Review<sup>2</sup> which gives detailed guidance on reporting requirements for studies with a favourable opinion, including, notifying substantial amendments, notification of serious breaches of the protocol, progress reports and notifying SFEC of the end of the study.

### **Feedback**

You are invited to give your view of the service that you have received from the Faculty Ethics Committee. If you wish to make your views known please contact the administrator at [ethics-sci@port.ac.uk](mailto:ethics-sci@port.ac.uk)

Appendix 4. SFEC 2016-061C Favourable Opinion, Notice of Substantial Amendment, Study 4&5 (Chapters 4&5)



Eglantine Julle-Daniere  
Department of Psychology  
University of Portsmouth

[eglantine.julle-daniere@port.ac.uk](mailto:eglantine.julle-daniere@port.ac.uk)

**Science Faculty Ethics Committee**

Science Faculty Office  
University of Portsmouth  
St Michael's Building  
White Swan Road  
PORTSMOUTH  
PO1 2DT

023 9284 3379  
[ethics-sci@port.ac.uk](mailto:ethics-sci@port.ac.uk)

5 October 2018

**FAVOURABLE ETHICAL OPINION - NOTIFICATION OF SUBSTANTIAL AMENDMENT**

**Study Title:** The behavioural expression of guilt – Can a facial expression of guilt be reliably identified by observers?

**Reference Number:** SFEC 2016-061C

**Date Submitted:** 2 October 2018

Thank you for submitting your proposal amendment to the Science Faculty Ethics Committee (SFEC) for ethical review in accordance with current procedures.

I am pleased to inform you that SFEC was content to grant a favourable ethical opinion of this proposal amendment on the basis described in the submitted documents listed at Annex A, and subject to standard general conditions (*See Annex B*).

We are of the understanding that the 2 parts of this experiment which participants previously completed both, has now simply been split into 2 separate experiments, to which participants can complete only one, or both separately. Please do let us know if this is not the intention.

Please note that the favourable opinion of SFEC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including the University of Portsmouth or supervisor, prior to the start of the study.

Wishing you every success in your research

A handwritten signature in black ink that reads 'Jim House'.

Dr Jim House  
Chair, Science Faculty Ethics Committee

**Annexes**

- A - Documents reviewed
- B - After ethical review - Guidance for researchers

Appendix 5. SFEC 2017-107 Favourable Opinion, After resubmission – Study 3

(Chapter 6)



UNIVERSITY OF  
PORTSMOUTH

Eglantine Julle-Daniere  
Department of Psychology  
University of Portsmouth

[Eglantine.Julle-Daniere@port.ac.uk](mailto:Eglantine.Julle-Daniere@port.ac.uk)

**Science Faculty Ethics Committee**

Science Faculty Office  
University of Portsmouth  
St Michael's Building  
White Swan Road  
PORTSMOUTH  
PO1 2DT

023 9284 3379

[ethics-sci@port.ac.uk](mailto:ethics-sci@port.ac.uk)

21 November 2017

**FAVOURABLE ETHICAL OPINION – FOLLOWING RESUBMISSION**

**Study Title:** "The social experience of guilt – How does the feeling of guilt influence social interactions?"

**Reference Number:** SFEC 2017-107

**Date Resubmitted:** 13 November 2017

Thank you for resubmitting your application to the Science Faculty Ethics Committee (SFEC) for ethical review in accordance with current procedures, for making the requested changes following the first SFEC review, and for the clarifications provided.

I am pleased to inform you that SFEC was content to grant a favourable ethical opinion of the above research on the basis described in the submitted documents listed at Annex A, and subject to standard general conditions (*See Annex B*).

Please note that the favourable opinion of SFEC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including the University of Portsmouth or supervisor, prior to the start of the study.

Wishing you every success in your research

A handwritten signature in black ink, appearing to be 'P. Morris'.

Dr Paul Morris  
Vice Chair, Science Faculty Ethics Committee

**Annexes**

A - Documents reviewed

B - After ethical review - Guidance for researchers

**Information:**



Appendix 6. UPR 16

# FORM UPR16

## Research Ethics Review Checklist



Please include this completed form as an appendix to your thesis (see the Research Degrees Operational Handbook for more information)

<b>Postgraduate Research Student (PGRS) Information</b>		<b>Student ID:</b>	795227
<b>PGRS Name:</b>	JULLE-DANIERE Eglantine		
<b>Department:</b>	Psychology	<b>First Supervisor:</b>	Prof Bridget WALLER
<b>Start Date:</b> (or progression date for Prof Doc students)	01/10/2015		
<b>Study Mode and Route:</b>	Part-time <input type="checkbox"/>	MPhil <input type="checkbox"/>	MD <input type="checkbox"/>
	Full-time <input checked="" type="checkbox"/>	PhD <input checked="" type="checkbox"/>	Professional Doctorate <input type="checkbox"/>

<b>Title of Thesis:</b>	The expression, experience, and social consequences of guilt: A cross-cultural study
<b>Thesis Word Count:</b> (excluding ancillary data)	44 710

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University's Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study

Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

### UKRIO Finished Research Checklist:

(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: <http://www.ukrio.org/what-we-do/code-of-practice-for-research/>)

a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
b) Have all contributions to knowledge been acknowledged?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
c) Have you complied with all agreements relating to intellectual property, publication and authorship?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
e) Does your research comply with all legal, ethical, and contractual requirements?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>

### Candidate Statement:

I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)

<b>Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC):</b>	SFEC 2015-097; 2016-061; 2016-061C; 2017-107
---	--

If you have *not* submitted your work for ethical review, and/or you have answered 'No' to one or more of questions a) to e), please explain below why this is so:

N/A

<b>Signed (PGRS):</b>		<b>Date:</b> 30/08/19
-----------------------	--	-----------------------