

Studies in Emotion and Social Interaction

The Expression of Emotion

Philosophical, Psychological
and Legal Perspectives

EDITED BY CATHARINE ABELL
AND JOEL SMITH



The Expression of Emotion

The Expression of Emotion is a collection of cutting-edge essays on emotional expression written by leading philosophers, psychologists and legal theorists. It highlights areas of interdisciplinary research interest, including facial expression, expressive action and the role of both normativity and context in emotion perception. While philosophical discussion of emotional expression has addressed the nature of expression and its relation to action theory, psychological work on the topic has focused on the specific mechanisms underpinning different facial expressions and their recognition. Further, work in both legal and political theory has had much to say about the normative role of emotional expressions, but would benefit from greater engagement with both psychological and philosophical research. In combining philosophical, psychological and legal work on emotional expression, the present volume brings these distinct approaches into a productive conversation.

Catharine Abell is Senior Lecturer in Philosophy at the University of Manchester. Her main research interests are in aesthetics, including the expression of emotion in art, depiction, fiction, genre and the nature of art. She was co-investigator on the Arts and Humanities Research Council (AHRC) Knowledge of Emotion project (2012–2015).

Joel Smith is Lecturer in Philosophy at the University of Manchester. His main research interests are the philosophy of mind, the philosophy of psychology and phenomenology, especially self-consciousness and the perception of others. He was the principal investigator of the AHRC Knowledge of Emotion project (2012–2015), and is the author of *Experiencing Phenomenology* (2016).

STUDIES IN EMOTION AND SOCIAL INTERACTION

Second Series

Series Editors

Keith Oatley
University of Toronto

Antony S. R. Manstead
Cardiff University

Titles Published in the Second Series:

The Psychology of Facial Expression, edited by
James A. Russell and José Miguel Fernández-Dols

Emotions, the Social Bond, and Human Reality: Part/Whole Analysis, by
Thomas J. Scheff

Intersubjective Communication and Emotion in Early Ontogeny, edited by
Stein Bråten

The Social Context of Nonverbal Behavior, edited by
Pierre Philippot, Robert S. Feldman, and Erik J. Coats

Communicating Emotion: Social, Moral, and Cultural Processes, by
Sally Planalp

Emotions across Languages and Cultures: Diversity and Universals, by
Anna Wierzbicka

Feeling and Thinking: The Role of Affect in Social Cognition, edited by
Joseph P. Forgas

Metaphor and Emotion: Language, Culture, and Body in Human Feeling, by
Zoltán Kövecses

Gender and Emotion: Social Psychological Perspectives, edited by
Agneta H. Fischer

Causes and Consequences of Feelings, by
Leonard Berkowitz

(continued at end of the book)

The Expression of Emotion

Philosophical, Psychological and Legal Perspectives

Edited by

Catharine Abell

University of Manchester

Joel Smith

University of Manchester



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE

UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107111059

© Catharine Abell and Joel Smith 2016

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2016

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Names: Abell, Catharine, editor. | Smith, Joel, 1976– editor.

Title: The expression of emotion : philosophical, psychological and legal perspectives / edited by Catharine Abell, Joel Smith.

Description: New York : Cambridge University Press, 2016. |

Series: Studies in emotion and social interaction |

Includes bibliographical references and index.

Identifiers: LCCN 2016023640 | ISBN 9781107111059 (hardback) |

ISBN 9781107527799 (paperback)

Subjects: LCSH: Emotions. | Emotions – Philosophical aspects.

Classification: LCC BF511.E97 2016 | DDC 152.4–dc23

LC record available at <https://lcn.loc.gov/2016023640>

ISBN 978-1-107-11105-9 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet Web sites referred to in this publication and does not guarantee that any content on such Web sites is, or will remain, accurate or appropriate.

Contents

| | |
|--|----------|
| <i>Notes on Contributors</i> | page vii |
| <i>Acknowledgements</i> | xi |
| Introduction: Emotional Expression <i>Joel Smith and Catharine Abell</i> | 1 |
| Part I Philosophical Perspectives | |
| 1 Expressing, Showing and Representing <i>Mitchell Green</i> | 25 |
| 2 Emotions and Their Expressions <i>James Sias and Dorit Bar-On</i> | 46 |
| 3 Expressive Actions <i>Christopher Bennett</i> | 73 |
| 4 Emotional Expression, Commitment and Joint Value <i>Bennett W. Helm</i> | 95 |
| 5 Collective Emotion and the Function of Expressive Behaviour <i>Carolyn Price</i> | 115 |
| Part II Psychological Perspectives | |
| 6 Emotional Expression <i>Hyisung C. Hwang and David Matsumoto</i> | 137 |
| 7 A Sceptical Look at Faces as Emotion Signals <i>James A. Russell</i> | 157 |

vi *Contents*

| | | |
|---------------------------------------|--|-----|
| 8 | Cultural Specificities in the Transmission and Decoding of Facial Expressions of Emotion <i>Rachael E. Jack</i> | 173 |
| 9 | The Impact of Context on the Perception of Emotions <i>Ursula Hess and Shlomo Hareli</i> | 199 |
| 10 | The Truth Is Written All Over Your Face! Involuntary Aspects of Emotional Facial Expressions <i>Alysha Baker, Pamela J. Black and Stephen Porter</i> | 219 |
| Part III Legal Perspectives | | |
| 11 | Regulating the Expression of Remorse and the Building of Moral Communities <i>Richard Weisman</i> | 247 |
| 12 | Share Your Grief but Not Your Anger: Victims and the Expression of Emotion in Criminal Justice <i>Susan A. Bandes</i> | 263 |
| 13 | Performing Anger to Signal Injustice: The Expression of Anger in Victim Impact Statements <i>Sarah Sorial</i> | 287 |
| | <i>Index</i> | 311 |

Notes on Contributors

Catharine Abell is Senior Lecturer in Philosophy at the University of Manchester. Her main research interests are in aesthetics, including the expression of emotion in art, depiction, fiction, genre and the nature of art.

Alysha Baker is currently completing a doctoral degree in experimental forensic psychology in the Centre for the Advancement of Psychological Science and Law (CAPSL) at the University of British Columbia Okanagan. Alysha has conducted research and published in the areas of emotional deception, deception detection, credibility assessment and memory.

Susan A. Bandes is Centennial Distinguished Professor of Law at DePaul University. She has edited a book on the role of emotion in law, *The Passions of Law* (2000).

Dorit Bar-On is currently the 2015–2016 Fellow at the Wissenschaftskolleg in Berlin. She has recently joined the Philosophy Department at the University of Connecticut (Storrs), after teaching many years at the University of North Carolina, Chapel Hill. Her main research interests are in the philosophy of language and mind, epistemology and metaethics. She is the director of the Expression, Communication and Origins of Meaning (ECOM) Research Group.

Christopher Bennett is Reader in Philosophy at the University of Sheffield. He is the author of *The Apology Ritual: A Philosophical Theory of Punishment* as well as numerous articles, and is currently Chief Editor of the *Journal of Applied Philosophy*.

Pamela J. Black is currently completing a doctoral degree in clinical forensic psychology in the Centre for the Advancement of Psychological Science and Law (CAPSL) at the University of British Columbia Okanagan. Pamela's research interests include interpersonal assessment, deception detection, psychopathy and the Dark Triad of personality.

Mitchell Green is Professor of Philosophy at the University of Connecticut. He is author of *Self-Expression*, and co-author (with J. Williams) of *Moore's Paradox: New Essays on Belief, Rationality and the First Person*.

Shlomo Hareli is Professor in the Department of Business Administration at the University of Haifa. His main research area is the social perception of emotions.

Bennett W. Helm is Elijah E. Kresge Professor of Philosophy at Franklin and Marshall College. He has published extensively on the nature and normativity of the emotions and is the author of *Emotional Reason: Deliberation, Motivation, and the Nature of Value* (2001).

Ursula Hess is Professor in the Department of Psychology at the Humboldt University of Berlin. She is the author of more than a hundred articles and book chapters.

Hyisung C. Hwang is Adjunct Faculty at San Francisco State University. Her research interests are in emotion, nonverbal behaviours and culture. She has also co-authored numerous scientific articles and book chapters on non-verbal behaviour, facial expressions and culture.

Rachael E. Jack, PhD is a Lecturer in the Institute of Neuroscience and Psychology at the University of Glasgow. She won the American Psychological Association (APA) New Investigator 2013 award, and has received funding from the Economic and Social Research Council (ESRC) Future Research Leaders scheme, Open Research Area and British Academy.

David Matsumoto is Professor of Psychology and Director of the Culture and Emotion Research Laboratory at San Francisco State University. He has studied culture, emotion, social interaction and communication for more than thirty years.

Stephen Porter is currently a researcher, instructor and consultant in the area of psychology and law in the Centre for the Advancement of Psychological Science and Law (CAPSL) at the University of British Columbia Okanagan. He has published numerous scholarly articles on deception detection, emotional deception, psychopathy and violent behaviour, and forensic aspects of memory.

Carolyn Price is Senior Lecturer in Philosophy at the Open University. Her research concerns what emotions are, what they are about and the norms by which they can be evaluated – questions she has recently addressed in a short monograph *Emotion* (2015).

James A. Russell is Professor of Psychology and Director of the Emotion Development Lab at Boston College, and Editor of the journal *Emotion Review*. He has published extensively on the expression and recognition of emotion through faces, including children's understanding of emotion and cultural influences on emotion expression.

James Sias is currently Assistant Professor of Philosophy at Dickinson College. His main research interests are in metaethics and moral psychology.

Joel Smith is Lecturer in Philosophy at the University of Manchester. He is the author of *Experiencing Phenomenology* (2016).

Sarah Sorial is Senior Lecturer in the Philosophy Program in the Faculty of Law, Humanities and Arts at the University of Wollongong, NSW, Australia. Her research specialization is at the intersection of political philosophy and philosophy of law. She has published widely on topics including the limits of free speech, hate speech, the expression of emotion, deliberative democracy and rights theory. She is the author of *Sedition and the Advocacy of Violence: Free Speech and Counter-terrorism* (2011).

Richard Weisman is Professor Emeritus in the Law and Society Program at York University in Toronto, Ontario, Canada. He is the author of *Showing Remorse: Law and the Social Control of Emotion* (2014).

Acknowledgements

The editors would like to thank the Arts and Humanities Research Council for generously funding the project Knowledge of Emotion, of which this collection is a result. We would like to thank the many people involved in this project, and especially those who contributed chapters to this volume for their hard work and patience during the process of bringing it all together. Thanks are due, too, to Rob Knowles for his work proofing and formatting the final version. Finally, we would like to thank Peter Goldie, who was originally to have been an investigator on this project but who sadly died shortly before it began.

Introduction: Emotional Expression

Joel Smith and Catharine Abell

Theorists of emotion typically recognize a number of features common to them: emotions are intentional, being directed towards objects in one's environment (including oneself); emotions involve the evaluation or appraisal of those objects as possessing various positive or negative values; emotions feel a certain way, in that there is something it is like to undergo an emotional experience; and finally emotions are expressed, involving a readiness or disposition to move one's body in a number of ways. Emotional expression in its variety – the topic of this volume of essays – is a phenomenon with which we are intimately familiar. It is something that we experience, both in ourselves and others, on a daily basis. As Edith Stein wrote, somewhat poetically,

I blush for shame, I irately clench my fist, I angrily furrow my brow,
I groan with pain, am jubilant with joy [...] as I live through the feeling,
I feel it terminate in an expression or release expression out of itself.
(Stein [1917]1970: 51)

But is a phenomenological description such as this supported by the scientific study and philosophical analysis of emotional expression? What is it for something to be an emotional expression and how do such expressions relate to other aspects of human psychology and behaviour? A common thought is that emotional expressions serve to communicate the emotional state of the expresser; indeed, the facial expression of emotion is often taken as the paradigm case in which the psychological states of others are made manifest to us (see, e.g., McNeill 2012; Smith 2013). Is this common-sense picture correct? In what sense can emotional expressions be thought of as communicative and what is it that they communicate? Further, emotional expressions are naturally thought to be subject to certain norms: a particular facial

expression is required for an apology to be considered sincere, another when receiving a gift and so on. What, we may ask, is the role of such norms in guiding our emotional behaviour and how do they interact with the 'release' of emotional expression that Stein speaks of? These questions are amongst those pursued in the chapters of this volume and may be thought to fall under three broad headings: the nature of emotional expression, the communicative role of emotional expression and the normative significance of emotional expression.

Before setting out these issues in more detail, however, we should say something about the interdisciplinary nature of the volume. Although each individual chapter comes to the subject of emotional expression from the perspective of one discipline in particular – philosophy, psychology or legal theory – all draw connections between their own themes and those addressed in a variety of other chapters in the volume. This is important since, to date, research in these three disciplines has been insufficiently integrated. For example, the philosophical focus on the nature of expression – one, in part, motivated by the use of the concept of expression in accounts of ethical language (Gibbard 1990) and self-awareness (Bar-On 2004) – has often relied on intuition rather than drawn on the extensive body of relevant psychological research. Psychological work, in its turn, although paying great attention to the specific mechanisms underpinning different facial expressions, has sometimes lacked an appreciation of what it is that makes something an expression of emotion at all. Finally, work on the topic within legal and political contexts would benefit from a more substantial engagement with both philosophical and psychological work on the nature and communicative role of emotional expression.

In the remainder of this introduction, therefore, not only do we expand on the three themes – the nature of emotion, the communicative role of emotion and the normative significance of emotion – we also draw connections between discussions of these issues within our three disciplines.

1.1 The Nature of Emotional Expression

Emotional expressions can be divided into expressions of emotion, on the one hand, and behaviour that is merely expressive of emotion, on the other. This is a distinction familiar from related discussions in aesthetics (Bennett, this volume; Davies 1994: Ch. 4; Hospers 1954–1955). As we use the term, some piece of a subject's behaviour can be an expression

of emotion only if there is in fact some emotion that the subject is in and that bears an appropriate relation to the expression. A piece of behaviour that is merely expressive of emotion need not meet this condition. Thus, someone may simulate joy by smiling. This may be expressive of joy, but it is not an expression of joy. So, although all emotional expressions are expressive of emotion, not all are expressions of emotion, since some are merely expressive. Of course, not all expression is emotional expression at all. Our utterances express our opinions, for example (see Green 2007 for an account of expression that generalises to a wide variety of cases). Our primary concern in this introduction, however, and the primary concern of the various contributors to this volume, is with emotional expressions that are expressions of emotion.

Many different forms of behaviour can be thought of as expressions of emotion. At one end of the spectrum we have behaviour intuitively characterised as voluntary action that expresses emotion (Bennett, Helm, Price, all in this volume). Paradigm cases of such actions would include jumping for joy or slamming a door in rage. Actions that express emotions – what Hursthouse (1991) calls ‘arational actions’ – have been the focus of much of the philosophical work on emotional expression for the reason that they have seemed difficult to fit into a widely accepted account of action explanation. On this account, actions are explained by belief–desire pairs (Davidson 1963). Annie opens the door because she wants to leave the house and believes that opening the door is a way to achieve that; Bob turns left because he wants to visit the supermarket and he believes that it is to his left and so on. Actions that express emotions present a puzzle for this view (one discussed in Bennett’s chapter in this volume), since they seem to be actions – angrily slamming the door is, after all, something that I choose to do and for which I may be held accountable – yet not easily explained in such a manner. It is not at all clear what belief–desire pair could explain Caroline’s jumping for joy, since it seems not to be a means to any end.

At the other end of the spectrum we have seemingly involuntary facial expressions of emotion, for example, wrinkling one’s nose at a disgusting smell, or smiling at a joke (see Baker, Black and Porter, this volume; Hwang and Matsumoto, this volume). It is with the analysis of the face and facial expressions of emotion that most psychological work in the area has been concerned. Stimulated by Darwin’s ([1872]2009) ground-breaking study, there was an explosion of psychological research on facial expression in the latter half of the twentieth century. Much of this research has been undertaken by proponents of a family

of views with roots in work by Tomkins (1962), Izard (1971) and Ekman and Friesen (1975). Broadly sketched, and ignoring disputes between proponents, this picture involves the postulation of a small number of discrete 'basic' emotions, typically including happiness, sadness, anger, disgust, fear, surprise and contempt. Each of these is associated with a characteristic facial expression which is automatically triggered as an individual enters into the emotional state (see Hwang and Matsumoto, this volume).

Lying between the cases of action and facial expression are a range of behaviours that are also intuitively thought to express emotion. For example, emotion is expressed not just through the face but through the whole body and done so in a way that need not involve action (Matsumoto and Hwang 2013; Wallbott 1998). Like facial expression, such bodily expression may be largely involuntary. A slumped posture, for example, may express grief, while a relaxed demeanour may express happiness. In addition, not only is emotion expressed in the movements of the body (face or otherwise), it is expressed also in a wide variety of vocalisations: from a growl of anger or exclamation of surprise, all the way to the subtle tone of voice that tells us that an utterance of 'Fine' means anything but (Frank, Maroulis and Griffin 2013; Kappas, Hess and Scherer 1991). Such vocal expressions of emotion, it would seem, span the range from involuntary to voluntary.

In each case, we can inquire into the role of biology, culture and context in producing both these emotional expressions and observers' subsequent judgements as to the emotional state of the expressing subject. For example, proponents of the view of facial expression that we sketched earlier typically argue that the basic emotions and associated facial expressions are a universal feature of human psychology (Ekman 1980; Hwang and Matsumoto, this volume). They thus form part of a biological core of human nature. This broad picture of emotion and emotional expression has, of course, been challenged in a number of ways and from a number of directions (see various chapters in Russell and Fernández-Dols 1997). For example, the claim that there is a small number of discrete basic emotions is questioned by proponents of dimensional approaches (Feldman Barrett 2006; Russell 2003, and this volume); the claim that the onset of emotion automatically triggers the associated facial expression has been challenged by naturalistic studies (Fernández-Dols and Crivelli 2013); the claim to universality has been challenged on a variety of grounds (Jack, this volume; Russell, this volume); and even the claim

that emotional expressions are typically caused by emotions has been challenged from the direction of ethology (Fridlund 1994).

That emotional expression appears to admit significant biological and cultural diversity raises the issue of the unity of the category of emotional expression. What features do these seemingly diverse phenomena share? Aside from the concern with action, mentioned earlier, philosophers working in this area have been keen to give an account of the nature of emotional expression (see Bennett, Green, Helm, Price, and Sias and Bar-On, all in this volume). An initial attempt to describe what is aimed for here might be to say that we want an account of what all and only emotional expressions have in common. But philosophical theories typically aspire to more than this, demanding not only an account of what emotional expressions in fact have in common, but also an account of what they must have in common. Only such an account, it might be thought, would provide us with the essence of emotional expression; with what it is that makes something an emotional expression as opposed, for example, to mere behaviour caused by emotion. For, although most would agree that a piece of behaviour's being caused by emotion is a necessary condition of being an emotional expression; few would consider this to be a sufficient condition. Anger may cause someone to bite their tongue but few would consider this to be an expression of anger.

Different accounts of the nature of emotional expression will have different consequences regarding whether difficult cases count as expressions of emotion because they will draw the line differently between expressions of emotion and mere behaviour caused by emotion. A case in point is blushing which, although classified as an expression of embarrassment by some views, is denied that status by others. Davis's (1988) account in terms of indication, according to which expressions must be sufficient to justify observers in believing that the subject is in the psychological state in question, appears to count blushes as genuine expressions. However, according to Green's (2007) signalling account, expressions must be designed to convey the information that the subject is in the relevant psychological state, a condition which he suggests blushing does not meet (2007: Ch. 2).

1.2 The Communicative Role of Emotional Expression

The two conceptions of expression just sketched tie the nature of emotional expression to the communication of information to observers.

Indeed, the role of emotional expression in our awareness of others' psychological states is something that has occupied both philosophers and psychologists in recent years. Accounts of the mechanisms underpinning, and the rational grounds for, our judgements about others' mentality sometimes appeal to emotional expression as making the psychological states of others manifest to observers (McNeill 2012; Smith 2013). Unlike our judgements about what other people believe or intend – which are typically considered to involve inferences based on either the application of a tacitly held theory of mind (Segal 1995) or the off-line simulation of their mental life (Goldman 2006) – it is sometimes claimed that our judgements about others' emotional states are grounded in perception with no need for such theorising or simulating (see Green, this volume; Sias and Bar-On, this volume). Emotional expressions, to adapt the familiar phrase, seem to be a window to the soul.

These deliverances of common sense tally with the basic emotions approach sketched earlier, according to which observers are typically able to recognize emotional expressions precisely as expressions of the basic emotions regardless of cultural differences. Indeed, psychologists working on emotion recognition often speak of it as a perceptual process (e.g., Feldman Barrett, Lindquist and Gendron 2007; Hess, Adams and Kleck 2009; Hess and Hareli, this volume).

Accounts of expression as involving the communication of emotion must say what notion of communication they employ. There are, broadly speaking, three varieties that could be distinguished. The first, which can quickly be put aside, thinks of communication as the mere transmission and receipt of information. It is clear that emotional expressions transmit a great deal of information that a suitably endowed observer will be able to pick up on. Much of it, however, will be irrelevant to our concerns. Smiling, for example, transmits information about the shape and colouration of one's teeth, but this is of little interest to theorists of emotion and emotional expression. It seems that a more restrictive conception of communication is required.

One such conception construes communication as the intentional transmission of information (in some reasonably liberal sense of 'intentional'). Linguistic communication, by and large, fits this model. Thus, an utterance of 'I am happy' intentionally communicates, to those within earshot, the information that one is happy. It might be thought that some emotional expression is of this sort. For example, one might pull a 'sad face' in order to let others know that one is disappointed.

Plausibly, however, the majority of emotional expressions will not count as communicative in this sense for the reason that they are involuntary. Intentionally doing something is surely a way of voluntarily doing it (though the converse is not true; see Anscombe 1957: §49). Spontaneous and involuntary emotional expressions are not realistically thought of as intended. The model, then, is really suitable only for those emotional expressions that are actions. More than this, some philosophers will reject the idea that there can be such a thing as an intended emotional expression, arguing that a necessary condition on being an emotional expression proper is that it not be done in order to achieve some further end, including that of communication (e.g., Goldie 2000: Ch. 5). On this conception of communication, then, at least the majority of emotional expressions will fail to be communicative.

An alternative way of thinking about communication is in terms of behaviour whose function it is to transmit information. Following Green (2007), something can have a function, *x*, either because it was designed by evolution, or designed by culture, to do *x*. This distinction might be thought of as one between natural and conventional functions. So, the markings on poisonous frogs have been designed by evolutionary processes to communicate danger, whereas the red flashing lights on various road systems have been culturally designed for a similar purpose. This provides a conception of communication that is, at least potentially, sufficiently flexible to accommodate all emotional expressions. It may be that some emotional expressions have been designed by natural selection pressures to communicate the information that one is in certain psychological states, for example, the smile to communicate happiness. On the other hand, perhaps some emotional expressions have been designed by the workings of particular cultures to communicate psychological states, for example the eye-roll to communicate exasperation or contempt.

This conception of emotional communication encompasses the previous one, since all behaviour that qualifies as the intentional communication of emotion thereby has the function of communicating emotion. But it is not limited to intentional communication, since functions may be designed by natural or cultural forces beyond the ken of the individual. On this construal, it becomes a largely empirical question whether expressive behaviour is communicative and, if so, exactly what is communicated.

The most natural response to these questions is that, yes, emotional expressions are communicative and that they communicate the

emotions they express. Such a view would follow, for example, from the claim that emotional expressions, although perhaps evolving for some other purpose, have subsequently been co-opted into the service of such a communicative role, perhaps being modified along the way so as to better perform this function (Shariff and Tracy 2011). On this popular view, one arguably derived from Darwin, the basic emotions are what are expressed and communicated by the characteristic facial expressions with which common sense associates them.

Such a position, although popular, is not uncontroversial. Proponents of the basic emotions approach claim that the basic emotions involuntarily cause the associated facial expression (Baker, Black and Porter, this volume; Hwang and Matsumoto, this volume). But this has been challenged. Fridlund (1994), for example, argues that evolutionary pressures would not give rise to such a 'leakage' of information about emotion. Rather, he claims, emotional expressions (or, to avoid the impression of begging the question, 'emotionally expressive facial behaviour') do not express emotion at all. They do perform a communicative function, but their role is to communicate the preferences and intentions of the subject in question. Thus, the stereotypical 'anger face' is not typically caused by anger, nor does it communicate that information; rather it serves as a threat, communicating an intention to attack; the stereotypical 'sad face' does not communicate sadness; rather it indicates to observers that one seeks to be comforted and so on.

Another view of the communicative function of emotional expressions – one that looks ahead to the discussion of norms in Section I.3 of this introduction – is that they communicate one's evaluative stance. It is typically recognized that emotion stands in close relation to evaluation (Helm 2001). This is most clearly brought out by appraisal theories (and those views that identify emotion with evaluative judgement [e.g., Nussbaum 2001]), according to which emotions result from the evaluative appraisal of a situation. Simplified, such a view might hold that anger results from the appraisal of an event as offensive, joy from the appraisal of it as pleasing and so on (Lazarus 1991). On such a picture, one feels emotion towards things that one cares about, and one's emotional reactions are indicative of the evaluative stance that one takes towards the objects of care (Helm 2001; Price 2013). As such, it may be that at least some expressions of emotion are performed precisely in order to do justice to the evaluative landscape of the situation (Bennett, this volume), and perhaps have the function of

communicating one's emotional stance both to those that share it and to those who do not (Helm, Price and Sorial, all in this volume).

The positions sketched here are not necessarily in conflict with one another, since it may be that emotional expressions serve to communicate more than a single piece of information. There is no doubt, however, that the first view has received the most attention in the philosophical literature. As mentioned earlier, a common thought is that emotional expressions provide a distinctive way of learning of the emotional states of others. An interesting and relatively unexplored question is the extent to which this common-sense idea is affected by the different conceptions of the communicative role of emotional expression mentioned earlier.

Another question is the extent to which cultural and contextual factors affect emotional recognition. For example, Jack (2013, this volume) presents evidence against the stronger claims to universality to be found within the basic emotions approach. Does this mean that expression is a distinctive route to the knowledge of others' emotions only between members of the same cultural group? The importance of this question is magnified when we consider contexts such as the courts in which much can ride on whether judges and juries read defendants' and others' facial expressions correctly.

I.3 The Normative Significance of Emotional Expression

There are, of course, differences in the ways that individuals and cultures express emotions. According to the basic emotion approach, these are to be explained by culturally various 'display rules' or 'feeling rules' (Hwang and Matsumoto, this volume; Weisman, this volume): norms that determine when a given expression is or is not appropriate. Our third guiding question concerns the norms that govern the expression of emotion, and the relation that such norms bear to both the nature and communicative role of emotional expression.

The interest in emotional expression within legal theory and forensic psychology has tended to focus on the issues of reliability and normative import (Bandes 2014; Bennett 2012; Black et al., 2012; Proeve and Tudor 2010). Courtrooms and other judicial settings represent very particular and highly formal contexts in which emotional expressions play an important role, not only in the context of fact-finding but also the more contested terrain of victim impact statements (Bandes, this volume; Sorial, this volume). In criminal trials, stakes are high and there is

often some motivation for defendants, victims and witnesses to appear to conform to prevailing norms. One way of doing this, of course, is either by suppressing or mimicking the relevant emotional expression (see Baker, Black and Porter, this volume). Notable examples would be the suppression of expressions of anger by victims for the reason that anger can sometimes turn juries and others against them (Bandes, this volume; Sorial, this volume), or the mimicking of expressions of remorse during probation hearings and other legal and extra-legal contexts (Weisman, this volume).

As these cases indicate, not only are there norms governing emotional expression in legal contexts, but also expectations concerning the consequences of expression: for example, catharsis for victims or apology for defendants (Bandes, this volume). The case of remorse is particularly striking since its expression can mean the difference between life and death, for the reason that (perhaps only implicitly) the courts acknowledge that the expression of remorse changes the normative landscape, making harsher punishment less appropriate (Maslen 2015; Proeve and Tudor 2010; Weisman 2014).

These cases also show the significance of reliability. Judges and juries are regularly in the position of evaluating the sincerity of others' emotional expressions. Since dissimulation is possible, and the stakes often high, it is crucial to know how to spot fakers of both masking and mimicking varieties (Baker, Black and Porter, this volume). Here it is important to have a clear sense both of exactly what is communicated via emotional expression and how the recognition of emotion works in dissembling and non-dissembling cases alike. Therefore, debates within psychology concerning what and how emotional expressions communicate have an obvious impact on the philosophical question of the capacity for the observation of emotional expression to ground our knowledge of others' emotional states. This, in turn, has significance for the legitimacy of relying on emotional expressions as evidence in high-stakes legal contexts.

As hinted in the previous section, emotional expression also arguably functions as a way of communicating values and thereby consolidating group identity. The expression of anger or disapproval, for example, can act to make clear the norms concerning behaviour within particular social contexts, as can the above-mentioned expressions of remorse serve to pave the way to re-admittance into a group after wrongdoing (see Helm, this volume, on 'communities of respect', and Weisman, this volume, on the moral community).

I.4 The Chapters in Context

Common sense tells us that emotional states are expressed in overt behaviour. A smile, a frown and a downcast gaze are each characteristic expressions of distinct emotional states, making those states easily recognisable in others. Thus, we might suppose, emotional expressions communicate emotion. Further, each sort of emotional expression is governed by certain familiar norms: wrongdoers should express remorse, expressions of anger are inappropriate in certain contexts and so on. The purpose of this volume is to evaluate these apparent platitudes concerning the nature and role of emotional expressions, either defending and elaborating or otherwise revising them. The volume does this by presenting together work on emotional expression written by philosophers, psychologists and legal theorists, with the subsidiary aim of highlighting areas of interdisciplinary interest. We have identified three central questions: What is emotional expression? What is the communicative role of emotional expression? What is the normative significance of emotional expression? Although each contribution to this volume pursues its own agenda from its own distinct perspective, each can be seen as making a contribution to answering one or more of these questions in order to gain a proper understanding of the nuances of emotional expression and its significance.

I.5 The Chapters

In his chapter, Mitchell Green considers the relation between expressive acts and artefacts, and representational acts and artefacts. There is a tendency to assume that the expressive capacities of acts and artefacts are wholly distinct from their representational capacities. However, drawing on his account of expression, Green argues that this is not the case. He begins by defending the view that it is possible to render at least some of one's emotions perceptible. If this is so, he argues, we can learn the contents of others' minds in ways other than by simulating them or theorising about them. He then goes on to argue that speech acts, such as apologies and assertions, extend the scope of those psychological states we can express beyond those we can make perceptible, because they are high-cost signals that ensure that speakers possess the psychological states they commit themselves to having by performing such acts. He argues that slur words also express speaker's attitudes towards the things those words represent. Although they do not

differ in representational content from their non-slurring counterparts, slur words conventionally express derogatory attitudes to their objects. Finally, he argues that metaphors enable speakers to express affective states by showing others how those states feel. Emotions and moods, he claims, lead us to perceive the world as being a certain way (whether or not it is, and whether or not we believe that it is). Metaphors enable us to convey how we perceive the world as being, and thus enable others to understand how we are construing things, and thereby how we feel. On Green's view, the power of speech acts, slurs and metaphors is due to the interrelatedness of their representational and expressive capacities.

Like Green, James Sias and Dorit Bar-On address the issue of whether or not we can literally perceive the emotions of others, as opposed to inferring their presence from people's behaviour. Their aim is not directly to defend the view that we can perceive emotions, but rather to determine what kinds of things emotions would have to be if that view were correct. They take issue with Green's claim that people's expressive behaviour enables us to perceive their emotions in virtue of providing us with evidence from which we infer (albeit nonconsciously, automatically and subpersonally) that they have those emotions. Green's use of the notion of inference, they claim, requires him to give up all that is appealing about the common-sense view that we can see a person's happiness in her face. They survey a range of philosophical accounts of emotion, and conclude that none can do justice to the claim that expressive behaviour can render emotions perceptible. What is needed, they argue, is an alternative view, according to which bodily changes are among the constitutive aspects of an emotion. Bodily changes, such as flushes, may at least sometimes have external manifestations that render them perceptible. They propose an account of emotions as complex states consisting of concerns, construals and bodily changes. On this view, they argue, emotional states are not subjective, private mental states as traditionally conceived. They are not states 'inside' individuals, but are instead states that whole individuals are in. Emotions have bodily changes as their parts. Thus, they claim, when we observe an emoting subject's facial expression and other expressive behaviours, we do not infer the presence of some 'inner' state of emotion; rather, we literally perceive the emotional state the subject is in, since to perceive a constitutive part or aspect of it is to perceive it.

Christopher Bennett addresses a philosophical problem that expressive actions pose to the rational explanation of action. Expressive acts

such as jumping for joy, while done intentionally, seem not to be done in order to achieve any particular purpose of the agent, and so seem to resist explanation by appeal to any such purpose. However, Bennett contends that such acts can indeed be explained by appeal to agents' purposes. They are performed intentionally for their own sake, he claims, as a constitutive part of doing justice to one's sense of a situation, or giving adequate external form to one's sense of that situation. Although they are not done for a purpose in the sense of comprising contingently effective means to a further end, they are purposive in the sense that they, as opposed to various other actions the agent might have performed, are appropriate vehicles for the emotions they express. Emotions, he argues, present situations to an agent in a distinctive way, leading the agent to see them under a certain characterisation. Consequently, emotions bring with them a sense of the way in which the situations at which they are directed matter. Expressive acts are appropriate vehicles for an emotion when they capture or reflect or do justice to our sense of what matters about the relevant situation. They do this by serving as symbols of that situation. Although one acts intentionally in creating the symbol, one does so simply in order to reflect one's sense of the nature of the situation. Bennett argues that it is the lack of any further end to which one's act might serve as means that has led some to think that expressive acts are not amenable to rational explanation.

In his contribution, Bennett Helm argues for a particular construal of emotional expressions, particularly expressions of reactive attitudes such as gratitude, resentment, approbation and guilt. He advocates a commitment account, according to which they express not just that one has a particular type of emotion directed at some target, but also one's evaluative attitude towards the focus of that emotion: the way in which one cares about it, or the import it has to one. According to Helm, what distinguishes expressions of reactive attitudes from mere reports of their contents is that they are expressions of commitments to the values that underlie them. As such, he argues, they play an important role in instituting joint values among the members of groups. The reactive attitudes are distinctive as a class of emotions in being ways of praising or blaming each other and thus holding each other responsible for what we do. Expressions of reactive attitudes, he argues, communicate this praise or blame to others. More than this, they communicate the demand for a suitable response to that praise or blame, in the form of recognition respect: the recognition of our status as members of a

community of respect. The commitments that expressions of the reactive attitudes communicate are commitments to norms of the relevant community and to the well-being of its members. In expressing the reactive attitudes, therefore, we call others to account in light of shared norms, thereby not merely communicating a positive or negative evaluation of the other's actions but also expressing the underlying values as our values in a way that can preserve and reinforce this shared sense of value.

In her chapter, Carolyn Price considers the extent to which emotional phenomena – emotional actions and processes of reasoning and reflection that arise in the course of an emotional response – can be collective. She chooses to address this question, rather than that of whether emotions themselves can be collective, because she doubts that there can be any adequate answer to the question of what an emotion is. She argues for the existence of two kinds of collective emotional phenomena: collective emotional action and collective emotional cognition. We can acknowledge both, she claims, without positing collective psychological states. She sees the existence of both as dependent in part on our capacity for expressive behaviour. Cases of collective action are sometimes cases of collective emotional action, she argues, because some types of emotional response motivate people to act together by employing expressive behaviour to plan and coordinate their collective actions. Likewise, instances of collective emotional cognition essentially involve expressive behaviour. To the extent that a case of collective planning involves the expression of emotional desires and thoughts, as when a group angrily plans a protest, it will be a case of emotional collective planning. It follows from this, Price argues, that complex emotional responses, perhaps even emotions themselves, might sometimes include a collective stage. She advocates construing such responses as processes with particular functions, which are supposed to unfold in a particular way. Such processes might incorporate collective emotional action or collective emotional cognition as stages. The function of an individual's outrage might be to ensure that a wrong is redressed, and this might require not just that she play a part in redressing the wrong, but that a group collectively redresses that wrong. Consequently, Price argues, the group's collective action should be seen as a stage in the individual's complex emotional response of outrage.

The basic emotions form the focus of [Hyisung C. Hwang's and David Matsumoto's chapter](#). These emotions, they argue, are evolutionarily adaptive, biologically innate and universal across humans

and even nonhuman primates. They endorse Darwin's principle of serviceable associated habits, according to which facial expressions are the behavioural remnants of more complete behavioural responses: humans express anger, for example, by furrowing the brow and displaying their teeth because these actions are part of an attack response. They review a range of evidence in favour of this view, including evidence that people from different cultures recognize one another's facial expressions of emotion, even when those cultures are wholly isolated from one another, and evidence that members of different cultures spontaneously produce very similar facial expressions when experiencing the same emotions. They also discuss studies showing that congenitally blind individuals spontaneously produce the same facial expressions of emotion as sighted individuals. Such studies demonstrate that the universality of facial expressions of the basic emotions has a biological base, they argue, by showing that the uniformity of facial expressions of emotion is not the result of learned imitation. Facial expressions of emotion are adaptive, they claim, because our spontaneous physiological reactions make it possible to manage survival risks immediately and to help ready ourselves for action. They also signal our emotions to others, providing them with crucial information concerning our emotions, intentions and attitudes. Nevertheless, they acknowledge that there is some cultural variability in facial expressions of emotion. This is the result, they claim, of different cultures imposing different display rules, which serve to maintain social order and facilitate cultural cohesion.

In his contribution, James Russell argues against the view of the emotions advocated by Hwang and Matsumoto, as evolutionarily adaptive, biologically innate and universal. This view (which Russell labels Basic Emotion Theory, or BET) holds that all human beings express the basic emotions using the same facial movements and that onlookers can read those emotions accurately from those expressions. However, Russell contends that the empirical evidence fails to support BET's claim that the facial expressions associated with the basic emotions are caused by, rather than merely correlated with, the basic emotions. He argues that non-emotional processes may instead be causally responsible for these facial movements and may explain the correlation between emotions and facial expressions. The evidence offered for BET, he contends, focuses less on how facial expressions are produced than on the alleged 'recognition of emotion' from them, and thus fails to establish the requisite causal link. Moreover, he identifies various flaws in the

design of the experiments usually cited in support of BET and cites a range of evidence that contradicts BET, appearing to show that emotion recognition varies with culture, language and education. Russell outlines an alternative account which he takes to fit the empirical evidence better than BET. This view, psychological construction, proposes that emotional episodes are not prefabricated, but are constructed in the moment to suit current circumstances. Episodes described as emotional consist of changes in various component processes, which are not themselves emotions, such that no clear boundary separates emotional from non-emotional episodes. This view, he argues, implies that the facial changes BET attributes to the basic emotions can be accounted for without invoking emotion concepts.

Like Russell, Rachael Jack is concerned to cast doubt on the claim that expressions of the basic emotions are universal. Although their universality appears to be supported by several decades of cross-cultural recognition studies, she argues that there is in fact marked cultural difference in recognition performance, questioning not just the universality hypothesis, but also the ability of the specific facial expression patterns used in recognition studies to communicate emotions across cultures. Indeed, she argues, traditional studies typically selected specific static facial expressions based on theoretical knowledge or observations and used above-chance performance to support claims of universal recognition, with the effect of masking any cultural differences in recognition performance. Moreover, researchers failed to explore the potential of other facial expressions to accurately communicate basic emotions outside of Western culture. She argues that there is a range of other facial expressions, both static and dynamic, that communicate basic emotions and reveal significant cross-cultural variability in expression. In support of this claim, she presents converging evidence from studies that use a novel psychophysical, data-driven approach. These studies show that, in facial recognition tasks, East Asians tend to focus attention on the eyes, whereas Westerners focus attention on both eyes and mouth. She argues that further studies, in which both static and dynamic facial expression patterns are randomly generated and the cultural receiver is asked to identify those that correspond to a given emotion, reveal that these differences in the way facial information is sampled and decoded are reflected in cross-cultural differences in how the face communicates emotion, with the eyes playing a greater role in conveying emotion amongst East Asians, while the mouth playing a more significant role amongst Westerners.

Ursula Hess and Shlomo Hareli discuss the role that context plays in the interpretation of facial expressions of emotion. Unlike Hwang and Matsumoto, who hold a version of basic emotion theory, and Russell, who advocates a constructivist account, Hess and Hareli endorse an appraisal theory of emotion, according to which emotions are elicited and differentiated through a series of appraisals of stimulus events based on the perceived nature of the event. The outcome of these appraisals is partially determined by the personality of the individual as well as their personal values and beliefs. Specific emotions are differentiated by the pattern of appraisal from which they result. Individuals' expressions of their emotions are often ambiguous, with the result that, in order to interpret those expressions, perceivers must rely on information about the situation in which the expression occurred, including information about the expresser, and on their own naïve emotion theories, including stereotypes and social and cultural encoding rules. Consequently, they argue, the same nonverbal behaviour when shown by a different person or in a different situation will be interpreted differently. Moreover, the interpreter's own personality, goals, motives and emotional states will influence how that behaviour is interpreted. Nevertheless, they argue, the pervasive influence of context on emotion perception does not mean that facial expressions fail to provide any information about expressers' emotional states. Rather, they claim, the role that context may play in interpretation is limited by the framework provided by the core appraisals constituting the emotion expressed. They conclude that perceivers' interpretations of facial expressions of emotion are influenced by contextual information only to the degree that this information is relevant within the appraisal pattern characterising the relevant emotion.

In their contribution to this volume, Alysha Baker, Pamela Black and Stephen Porter examine the issue of how insincere emotional expressions differ from sincere expressions of emotion and ask to what extent it is possible for observers to distinguish successfully between the two. Focussing on the basic emotions, they distinguish three different ways in which emotional facial expressions may be intentionally manipulated to produce insincere emotional expressions: by simulating (adopting an emotional expression in the absence of any real emotion), masking (concealing a felt emotion with the expression of a different, unfelt emotion) and neutralising (concealing a felt emotion with a neutral facial expression). They argue that psychopaths and those with high levels of emotional intelligence are better able than others to mask their

emotions and to produce insincere yet convincing emotional expressions. Nevertheless, they argue, intentional manipulation cannot produce emotional facial expressions that are indistinguishable from genuine expressions of emotion, both because not all of the muscles involved in the expression of basic emotions are under our voluntary control, and because we cannot completely prevent our powerfully felt actual emotions from registering on our faces. They conclude by considering the implications of these findings for high-stakes legal contexts, in which judgements as to the sincerity of emotional expressions are often critical to decisions regarding reduced sentences and early release.

In his contribution, Richard Weisman discusses the regulative role of the law in determining when expressions of remorse are appropriate, and what form those expressions should take. The law, he notes, is one of many sources of such social and moral regulation, the others including the various extra-legal moral communities to which individuals may belong. However, he argues, because the law is public and carries with it the authority of the state, the pronouncements of a court that a wrongdoer did not show remorse when they should have or that their expressions of remorse were credible or not credible has an impact on the community as a whole. Such pronouncements may serve to either reinforce or to contradict the norms of the moral community to which wrongdoers belong, with the result that, in cases of contradiction, wrongdoers must either act in conformity with the expectations of the moral community sanctioned by the court, thereby alienating themselves from their own moral communities, or act in conformity with the expectations of their own moral communities, thereby forgoing the benefits of mitigation and risking a more severe sentence. When wrongdoers fail to act in conformity with the court's norms, this is a cause for collective moral outrage. Consequently, Weisman argues, it is not just the transgressive act that plays a role in legal decision making, but also the transgressor's feelings about that act and the manner in which those feelings are shown. The role of moral communities in regulating expressions of remorse is particularly evident, he argues, during periods of social and individual change, when moral boundaries are liable to shift and individual feelings of remorse and expressions thereof are liable to change along with changes to the moral communities of which they seek membership.

[Susan Bandes's chapter](#) addresses the role that victims' expressions of emotion play in legal and, particularly, courtroom contexts. She argues that whether or not they have a legitimate role to play and what form

these expressions should take depends on what goals victims' expressions of emotion are intended to serve and on whether those goals are practically achievable or normatively desirable. She argues that victim impact statements and restorative justice conferences, both of which explicitly recognize a role for victims' expressions of emotion, have goals which are often poorly articulated, poorly understood and ambiguous. These goals include providing information to judges and juries, promoting victim catharsis and healing, eliciting remorse, empathy or shame in the defendant and facilitating forgiveness. The criminal justice system, she argues, affects how victims express their emotions, because it lays down implicit and explicit rules governing what emotions a victim may express. If victims' expressions of emotion fail to conform to these display rules, she argues, they are likely to be either misinterpreted or to elicit responses other than those sought. This is problematic, not just because it discourages the sincere expression of emotion, preventing the achievement of the aims of catharsis, healing and information provision, but also because the rules may be biased against certain types of victim. This may be because their interests are not aligned with those of the prosecution, or because they have gender, race or other characteristics that elicit different expectations regarding their expressions of emotion.

Sarah Sorial's contribution, like Bandes's, addresses the expression of emotion in victim impact statements. However, whereas Bandes is concerned to emphasise the ways in which the criminal justice system limits the effectiveness of victims' expressions of emotions in achieving the various goals of these statements, Sorial argues that victims' expressions of anger have an important and overlooked function in the context of victim impact statements. Drawing on J. L. Austin's important observation that words can be used to perform a variety of acts other than simply describing or stating, she advocates construing victims' expressions of anger as a kind of performance, which signals or communicates that an injustice has occurred which needs to be recognized and rectified. Such expressions, on this conception, are not mere descriptions of how victims feel, but rather ways of doing a variety of valuable things, such as securing institutional recognition of the injustice, facilitating empathy, eliciting the imposition of appropriate penalties and repairing community relations that have been damaged as a result of the injustice. The point of a victim impact statement, on this understanding, is not catharsis or healing for the victim, but recognition and redress for injustice. However, for victims' expressions of anger to have this effect, she argues, attention must be paid to how they

express their anger. In particular, victims' expressions should be proportionate to the moral issue at hand, and the language used to express victims' anger should be restrained. When these conditions are met, she claims, performative expressions of anger can enable better outcomes for victims.

References

- Anscombe, G. E. M. 1957. *Intention*, 2nd edn. Oxford: Blackwell.
- Bandes, S. 2014. 'Remorse, Demeanour, and the Consequences of Misinterpretation: The Limits of Law as a Window into the Soul', *Journal of Law, Religion, and State* 3: 170–199.
- Bar-On, D. 2004. *Speaking My Mind: Expression and Self-Knowledge*. Oxford: Oxford University Press.
- Bennett, C. 2012. 'Excuses, Justifications and the Normativity of Expressive Behaviour', *Oxford Journal of Legal Studies* 32: 563–581.
- Black, P., Porter, S. Baker, A. and Korva, N. 2012. 'Uncovering the Secrets of the Human Face: The Role of the Face in Pro-Social and Forensic Contexts', in S. E. Carter and V. T. Bailey (eds.) *Facial Expressions: Dynamic Patterns, Impairments, and Social Perceptions*. Hauppauge, NY: Nova Science Publishers.
- Darwin, C. [1872]2009. *The Expression of the Emotions in Man and Animals*, 4th edn. Oxford: Oxford University Press.
- Davidson, D. 1963. 'Actions, Reasons, and Causes'. Reprinted in D. Davidson, 2001, *Essays on Actions and Events*, 2nd Edition. Oxford: Clarendon Press.
- Davies, S. 1994. *Musical Meaning and Expression*. Ithaca, NY: Cornell University Press.
- Davis, W. 1988. 'Expression of Emotion', *American Philosophical Quarterly* 25: 279–291.
- Ekman, P. 1980. *The Face of Man: Expressions of Universal Emotions in a New Guinea Village*. New York: Garland STPM Press.
- Ekman, P. and Friesen, W. 1975. *Unmasking the Face*. Englewood Cliffs, NJ: Prentice-Hall.
- Feldman Barrett, L. 2006. 'Are Emotions Natural Kinds?' *Perspectives on Psychological Science* 1: 28–58.
- Feldman Barrett, L., Lindquist, K. and Gendron, M. 2007. 'Language as Context for Emotion Perception', *Trends in Cognitive Science* 11: 327–332.
- Ferdández-Dols, J. and Crivelli, C. 2013. 'Emotion and Expression: Naturalistic Studies', *Emotion Review* 5: 24–29.
- Frank, M. G., Maroulis, A. and Griffin, D. 2013. 'The Voice', in Matsumoto, Frank, and Hwang (eds.) *Nonverbal Communication: Science and Applications*. Thousand Oaks, CA: Sage, pp. 53–74.
- Fridlund, A. 1994. *Human Facial Expression: An Evolutionary View*. San Diego, CA: Academic Press.
- Gibbard, A. 1990. *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.
- Goldie, P. 2000. *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press.

- Goldman, A. 2006. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford: Oxford University Press.
- Green, M. 2007. *Self-Expression*. Oxford: Clarendon Press.
- Helm, B. W. 2001. *Emotional Reason: Deliberation, Motivation, and the Nature of Value*. Cambridge: Cambridge University Press.
- Hess, U., Adams, R. B. Jr. and Kleck, R. 2009. 'The Categorical Perception of Emotions and Traits', *Social Cognition* 27: 320–326.
- Hospers, J. 1954–1955. 'The Concept of Artistic Expression', *Proceedings of the Aristotelian Society, New Series* 55: 313–344.
- Hursthouse, R. 1991. 'Arational Actions', *Journal of Philosophy* 88: 57–68.
- Izard, C. 1971. *The Face of Emotion*. New York: Appleton Century Crofts.
- Jack, R. 2013. 'Culture and Facial Expressions of Emotion', *Visual Cognition* 21: 1248–1286.
- Kappas, A., Hess, U. and Scherer, K. R. 1991. 'Voice and Emotion', in Feldman and Rimé (eds.) *Fundamentals of Non-Verbal Behaviour*. Cambridge: Cambridge University Press.
- Lazarus, R. 1991. *Emotion and Adaptation*. New York: Oxford University Press.
- Maslen, H. 2015. *Remorse, Penal Theory, and Sentencing*. Oxford: Hart Publishing.
- Matsumoto, D. and Hwang, H. S. 2013. 'Body and Gestures', in Matsumoto, Frank, and Hwang (eds.) *Nonverbal Communication: Science and Applications*. Thousand Oaks, CA: Sage, pp. 75–96.
- McNeill, W. E. S. 2012. 'On Seeing That Someone Is Angry', *European Journal of Philosophy* 20: 575–597.
- Nussbaum, M. 2001. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.
- Price, C. 2013. 'The Problem of Emotional Significance', *Acta Analytica* 28: 189–206.
- Proeve, M. and Tudor, S. 2010. *Remorse: Psychological and Jurisprudential Perspectives*. Farnham, Surrey: Ashgate Press.
- Russell, J. 2003. 'Core Affect and the Psychological Construction of Emotion', *Psychological Review* 110: 145–172.
- Russell, J. and Fernández-Dols, J. (eds.) 1997. *The Psychology of Facial Expression*. Cambridge: Cambridge University Press.
- Segal, G. 1995. 'The Modularity of Theory of Mind', in Carruthers and Smith (eds.) *Theories of Theories of Mind*. Cambridge: Cambridge University Press, pp. 141–157.
- Shariff, A. and Tracy, J. 2011. 'What Are Emotion Expressions for?', *Current Directions in Psychological Science* 20: 395–399.
- Smith, J. 2013. 'The Phenomenology of Face-to-Face Mindreading', *Philosophy and Phenomenological Research* 87: 274–293.
- Stein, E. [1917]1970. *On the Problem of Empathy*. Translated by Waltrout Stein. The Hague: Martinus Nijhoff.
- Tomkins, S. 1962. *Affect, Imagery, Consciousness: Vol.1 The Positive Affects*. New York: Springer.
- Wallbott, H. G. 1998. 'Bodily Expression of Emotion', *European Journal of Social Psychology* 28: 879–896.
- Weisman, R. 2014. *Showing Remorse: Law and the Social Contract of Emotion*. Farnham, Surrey: Ashgate Press.

PART I

Philosophical Perspectives

1. Expressing, Showing and Representing

Mitchell Green

1.1 Expression as Genitive or Generic

Various forms of expressive behaviour are best seen as solutions to the problem of how to make our psychological states and processes known to others.¹ In some cases, natural selection appears to have solved this problem by making our psychological states perceptible to conspecifics.² In others we indicate those states in ways that, while not enabling perception thereof, still put appropriately situated others in a position to discern what state of mind we are in. True to their origins in the Latin concept *exprimere* (meaning *to press out*), many of the most significant modern practices of expression require those engaging in them to put something of themselves on the table: at least to express *one's* state of mind, one must be in that state, be it remorse, concern, resentment or certainty about the future. In this way, expression differs from the blander but more labile notion of representation.

Let us set aside semantic expression, such as occurs when an indicative sentence expresses a proposition or a predicate expresses a concept. Within the remainder of the phenomenon of expression, we may

¹ Earlier versions of this chapter were presented at the University of Manchester Workshop on Emotion and Expression, June, 2013; the Eastern Division of the American Philosophical Association, December, 2013; University of Tübingen, June, 2014; and the University of Connecticut, February, 2014. My thanks to audiences on those occasions, as well as to Dorit Bar-On, and to the editors of this volume for their insightful comments and suggestions for improvement.

² In what follows, for brevity I refer primarily to psychological states even though many mental items are more accurately described as processes. Many emotions, for instance, have characteristic diachronic features. Also, co-evolution appears to have enabled some between-species sensitivity to psychological states. Topál et al. (2009), for instance, provide evidence for such sensitivity to others' states of mind between dogs and humans.

differentiate between self-expression and expressiveness. An account of self-expression explains what it is to express one's own psychological state. By contrast, an account of expressiveness tells us what it is for an agent to engage in expressive behaviour even when she is not expressing her own state of mind or feeling; such an account also tells us how an inanimate object may have an affective dimension such as exuberance or melancholy. (We describe some paintings as exuberant and even speak of wind as melancholy.) Non-semantic expression encompasses both self-expression and expressiveness.

Green's *Self-Expression* (2007) does not offer a general account of non-semantic expression, but rather offers only accounts of self-expression and of certain forms of expressiveness. Yet the work has the resources to provide an account of non-semantic expression per se. (I restrict the theory to non-semantic cases because it would take us too far afield to offer an account of what it is for words, phrases and sentences to express propositions and other kinds of content.) Accordingly, a general account of non-semantic³ expression is as follows: *A expresses psychological state B* just in case A either expresses her B (expresses a psychological state that is her own), or A engages in behaviour or possesses features that are expressive of B. A might be an agent, but to satisfy the conditions of the second disjunct, A could also be an artefact or even a natural phenomenon. Let us call this the *genitive-or-generic account of expression*.

The account I offer of expression rests on the notions of signalling and showing. In signalling, one does something or has a characteristic that is designed to convey information. The notion of design in question may be that of evolution by natural selection, cultural evolution or intention. As a result, one might signal with no intention of doing so. For instance, on some views of the natural history of human facial expression, so-called basic emotions have pan-cultural facial signatures that evolved to communicate information about one's affective state.⁴ If this is correct, however, these facial behaviours communicate affective information involuntarily.

³ In what follows I will suppress this qualification but ask to be understood as assuming it. The general account of expression develops that suggested in Green (2011). Observe also that I am only characterizing the process of expressing psychological states. Some writers hold that it is also possible to express character traits such as courage and integrity. The present chapter remains neutral on the question.

⁴ We will return to consider challenges to this view of basic emotions in [Section 1.3](#) of this chapter.

In drawing on the notion of showing, I distinguish among three forms that it takes: showing-that, showing-a and showing-how. In showing-that, an agent or a phenomenon provides strong evidence for the obtaining of a state of affairs. Susan's astrophysical proof shows that Galaxy NR-439 is inhabited by a pulsar; the periodicity of the cicadas' chirping shows that it is hot outside; and the vagaries of the real estate market show that the nation's economy has yet to recover from the recent economic downturn. The second notion of showing that I employ is that of showing-a, in which a behaviour or other aspect of an entity makes an object perceptible. I might show you the lacerations on my neck, and in so doing enable you to see them. By playing a recording of its call I might show you the sound of a kestrel. Finally, one might show someone how to do something or how an emotion feels or what an experience is like. I could show someone how to build a fire with flint and twigs, what it feels like to be ostracized or how sulphur smells.⁵

How does one express one's own psychological state? One does this by engaging in behaviour (not necessarily voluntary) that both shows (in any of the three ways mentioned) and signals that state. By contrast, an act, phenomenon or artefact is expressive just in case it shows how a psychological state, not necessarily one's own, characteristically appears, behaves or feels. (The groaning of the trees in the gale shows us how human sadness sounds, and that is why the trees sound melancholy; the searing guitar riff sounds angry; and so on. We will return in [Section 1.6](#) to explore how a behaviour could show how a psychological state feels.)

The requirement of showing invoked in the accounts of both self-expression and expressiveness is substantial, and one aim of this chapter will be to motivate a view of expressing as being so demanding. Further, the notion of expression I suggest is broad enough to include behaviours outside of our conscious control, behaviours that are spontaneous even though they could have been suppressed, as well as those depending upon arduous training and premeditation. Thus while the most commonly discussed cases of expression might be scowling, screaming in pain or exuberantly hugging a loved one, we do well to keep in view forms of expression whose appreciation might require

⁵ I could show you how sulphur smells without obliging you to smell any sulphur: a scratch-and-sniff picture might chemically recreate the smell of sulphur without putting you in perceptual contact with the element. This is why showing-how does not imply showing-a.

intimate knowledge of a couple's past, or of the last century of developments in atonal music. So too an actor expresses emotion. He will also express his emotion if he is method acting. Otherwise, he will only behave expressively by expressing the emotions experienced by the characters he is portraying. In the movie *There Will Be Blood*, Daniel Day Lewis expresses the rage felt by the character Daniel Plainview when the preacher Eli Sunday obliges him to confess his iniquity. The actor is also expressing himself if he feels that rage and he does something that signals and shows it. Similarly, an attorney might express the desires of the deceased concerning how her estate shall be divided among her survivors. Although it is doubtful that one can act on behalf of a fictional character, one can evidently act on behalf of a person who has died. This suggests that the attorney can express desires on behalf of his deceased client. So long as we keep in view a distinction between self-expression and expressiveness, it would appear that one can express someone else's psychological state and not just one's own.

One challenge for any satisfying theory of expression is to provide an account both illuminating and broad enough to encompass the rich variety of expression-involving phenomena. To that end, we do well to be sceptical of a stark dichotomy between expression and representation.⁶ Instead, a representation (for instance, an indicative sentence, noun phrase or verb phrase) may have an expressive dimension, and using such a representational device in communication may add yet another layer. We see this in many speech acts, in much figurative language and in slurs. Accordingly, by way of assessing its merits, in what follows, I will apply the genitive-or-generic account of expression to these three cases. Before doing so, however, I turn to the issue of perceiving emotions.

1.2 Some Emotions Are Perceptible

It does not follow from the genitive-or-generic account of expression that it is possible to perceive an individual's emotions. It is compatible with the theory sketched so far that the only way in which we show our emotions is by showing that we are undergoing them or by showing how they feel. However, it would be a significant finding if we could

⁶ Bar-On's (2004) neo-expressivist theory of avowals, on which an avowal may also serve to express the state it avows, illustrates well the value of not taking expressing and representing as mutually exclusive activities.

establish that natural selection, or perhaps even cultural evolution, solved the problem of how we make our minds known to others by occasionally making their contents perceptible. What is more, the present account of expression provides the means to ask in a sharp way whether emotions may be perceived. For we might ask, when an agent expresses an emotion, is this ever because she has shown-a that emotion, thereby enabling an appropriately attuned observer to perceive what she is expressing?

An affirmative answer may go as follows.⁷ We can perceive a complex object by perceiving one of its characteristic components. That is why we perceive a durian by seeing one of its facing surfaces. It does not follow that in so doing the perceiver undergoes a transition from one perceptual experience to another. In such a case there may well be only one perceptual experience, namely, of the durian's facing surface. For all we know, a typical perceptual experience of a durian may be akin to Type I blindsight: one may show her awareness of the fruit's location in egocentric space, manoeuvre around it and otherwise act as if she has perceived it, without requiring a phenomenologically distinct experience from that of its facing surface.⁸

So, too, an emotion is a complex consisting of characteristic physiological, behavioural (including vocal and facial expression) and experiential components.⁹ A facial expression of emotion is thus a characteristic component of that emotion; likewise for vocal and other behavioural characteristics, such as gait and posture. If we perceive a fruit by perceiving one of its facing surfaces, then, I suggest, there is an equally good case that may be made for the conclusion that we perceive emotions by perceiving their 'facing surfaces' (although, see Russell, this volume, for criticisms of the view that such perceptible features are parts of emotions). We may qualify the conclusion to take account of the possibility of a pattern of separation between characteristic components and what they are components of. So just as we may only see everyday objects in contexts in which their facing surfaces normally are

⁷ Green (2007) suggests an approach along these lines, and develops it in Green (2010). See Bar-On and Sias, this volume, for an alternative perspective.

⁸ Weiskrantz (2009) provides a comprehensive discussion of Type I and Type II blindsight.

⁹ Griffiths (1998) provides a succinct overview of this approach. This approach does not settle disputes between, say, cognitivist and non-cognitivist theories, as both sides could accept it. For instance, a view of emotions as concern-based construals (a modified cognitivism) might explicate the notion of construal in experiential or physiological terms or a combination of the two.

attached to them, so too we may only perceive emotions in contexts in which their ‘facing surfaces’ are systematically attached: on an island full of actors, we may not perceive emotions by perceiving their facing surfaces. Similarly, as Stout (2010) observes, our perceptual experiences of emotional displays need not be fleeting, but instead may be (and probably most often are) extended over a period during which we observe another’s affective state take its course.

This conclusion about perception of emotional displays is also enhanced by acknowledging that seeing someone’s anger is more mediate than is, for example, seeing the redness of an apple. As Smith (2015) observes, the looking red of the apple is in a way basic: its looking red does not depend on some other more basic way that it looks. By contrast, Takeo’s looking angry does seem to depend on some other more basic ways that he looks: the slant of his eyebrows, the quavering mouth, clenched fists, his upturned head. Even if we do not attend to these more basic aspects of how Takeo looks, they would appear to be available to introspective awareness. In this respect, perception of emotion is phenomenologically distinct from perception of colours, shapes and the like. It is still true that we perceive Takeo’s anger, but we may now also account for how his looking angry is different from the way in which the red apple looks red.¹⁰

One might worry that if Takeo’s looking angry depends on some more basic ways that he looks, then his looking the way that he does will in some sense be the result of inference, and thus not a case of perception after all.¹¹ It is widely agreed that if you perceive a durian, it is not by virtue of consciously inferring its presence from a set of propositions entailing it. However, the notion of inference I am invoking is not the high road one commonly used in logic and epistemology, wherein inference is a conscious, or at least consciously accessible, process of

¹⁰ Jacob (2011) offers a dilemma that would seem to challenge the thesis that emotions may be perceived, even as that thesis has been here qualified. A person’s bodily manifestations of emotion, Jacob suggests, either constitute that emotion or they do not. If they do not, then it is difficult to see how by perceiving a bodily manifestation of emotion, one also perceives that emotion. By contrast, if those manifestations constitute emotion, then we seem to have reverted to a form of behaviourism. The current strategy for defending the view that emotions are perceptible, invoking as it does the idea that one can perceive a complex by perceiving one of its characteristic components, has the resources to respond to Jacob’s dilemma: it offers an account of how bodily manifestations, though distinct from emotions, nevertheless enable perception thereof.

¹¹ See, for instance, O’Shea (2012), particularly [note 26](#).

deducing conclusions from sets of propositional premises. Instead the notion I appeal to is the (low-road) automatic, encapsulated, typically unconscious process, or set of processes that enable us to compute sentence meaning, the spatial distance of perceived objects, colour constancy and the like.¹² After all, there must be an inferential process that enables you to interpret Hazlitt's elliptical remark, 'Prosperity is a great teacher, adversity a greater', as saying *inter alia* that adversity is a greater teacher than is prosperity. It is implausible that this process involves consciously accessible manipulation of propositions. And yet it is widely held that sentence understanding is at the very least perception-like.¹³ Furthermore, inferential processes, of either the high-or low-level kind, are in their nature subject to norms: where there is inference there must be the possibility of fallacy or an analogue thereof. Yet we do see such errors in processes I am describing as low-road inferential: perceptual illusions are just such cases, as are the misconstruals associated with garden-path sentences like 'The horse raced past the barn fell' and 'Fat people eat accumulates'.

The perception-of-emotion thesis suggests that we can learn the contents of others' minds in ways beside theorization or simulation, at least as the Theory and Simulation Theory are typically formulated. It also suggests that one may in some instances learn about one's own affective state by perceiving it, for instance, by looking in the mirror or listening to a recording of oneself speaking. The idea that in some cases we can perceive emotions comports, further, with common sense, and has lately gained the interest of some proponents of embodied cognition.¹⁴ It would, however, appear overly ambitious to conclude that whenever a person expresses an emotion she has made it perceptible. Too many emotions are too far removed from sensuous characteristics for this across-the-board claim to be tenable, even after we have taken the measure of Smith's distinction between basic and non-basic ways things look. I might express my disapproval of a

¹² Hatfield (2002) provides a historically detailed discussion that raises doubts about the prospects for a viable notion of unconscious inference. However, the conception of inference his discussion presupposes is needlessly narrow. For instance, Hatfield construes connectionist models for visual cognition as non-inferential alternatives to inferential accounts, whereas, so long as those models perform calculations, and those calculations can be assessed for their accuracy, they should be construed as inferential as well.

¹³ Smith (2009) discusses the matter in detail.

¹⁴ See, for instance, Gallagher (2008) and the many responses it has provoked, including Krueger (2012).

group's activities by being pointedly absent from their meetings. It is simply unclear what it would mean to say that in staying away from their meetings I have made my disapproval available to perception. So too, a diplomatically crafted thank you note might express your gratitude to a host for inviting you to their black tie event. Even if your expression of thanks is sincere, here too it is difficult to see the sense in which you've made your gratitude perceptible. But 'difficult to see the sense in which' does not mean 'false'. Perhaps in such cases one has made one's emotion perceptible. I doubt, however, that we are currently in a position to establish that one has, and as a result a more cautious policy is to construe such cases as depending on showing-that rather than showing-a.

It nevertheless seems to some that signalling and showing(-that) an affective state cannot be sufficient to express that state. Bar-On offers the example of, 'a perfectly empathetic mother hen who was able to produce knowledge-enabling signals of her chicks' hunger or fear' (2010: 217). By hypothesis what the mother hen produces are signals and not just indications, so they are designed to provide information. Further, these signals are sufficiently reliable that they show that the chicks are in a certain psychological state rather than just providing evidence that they are. Nevertheless, Bar-On denies that the mother hen is expressing the affective states of her offspring.

It is not clear what this example is supposed to challenge. The account given in [Section 1.1](#) of self-expression is defined only for those who are in the psychological state putatively expressed. As a result, to challenge the sufficiency of that definition one would need a mother hen who is hungry or fearful, who meets the further constraints of the definition, but who is intuitively not expressing her own hunger or fear. If we take Bar-On's epithet, 'perfectly empathetic' literally, then presumably the mother hen is hungry or fearful; but then it is not clear that we have a counterexample. By contrast, if we do not take the epithet literally, then the example could not be one that would challenge the definition of self-expression.¹⁵ Nonetheless, the genitive-or-generic theory leaves open the possibility of the mother hen expressing a fear or hunger not necessarily her own, so long as her behaviour is

¹⁵ Green (2008) argues that some forms of empathy require imagining oneself into another's emotional situation, but it is doubtful that the hen is capable of such a thing. Instead, if she is empathetic, it would seem on account of sharing the emotions of her offspring. But then, she would appear after all to be expressing her own emotions.

expressive of hunger or fear. For all that Bar-On's case tells us, the hen may be doing just this.¹⁶

1.3 If Emotions Are Not Natural Kinds

A widely accepted set of claims about emotions and their expressions construes them as being, or being very like, natural kinds: on this view, there is a small group of emotions (usually identified as anger, fear, surprise, happiness, disgust and sadness – there is some debate over whether contempt should be included among these) that are experienced at one time or another by nearly every member of our species; each of these emotions has a characteristic pattern of physiological, behavioural, facial-expressive and subjective/experiential markers. The facial-expressive marker associated with each of these basic emotions is readily recognized by our conspecifics as being an indication of the emotion of which it is the marker. (A version of this view is defended by Hwang and Matsumoto, this volume. For criticisms of it, see Russell and Jack, both this volume.¹⁷)

If the natural kinds view of emotions is correct, then one can readily discern how certain human emotions are complexes having certain characteristic components. Furthermore, some of these characteristic components are readily observable by normal conspecifics. That, in

¹⁶ As another challenge to the view that signalling and showing can ever be sufficient for emotional expression, Bar-On provides a case of compulsive self-reporters: creatures who compulsively issue reports of the mental states they are in, and are unable to lie about them. Bar-On asks, 'Would the creatures' compulsive self-reports be rendered any more instances of self-expression just in virtue of their unassailable reliability?' (2010: 217). My own accounts of expression, self-expression and expressiveness do not make commitment about ways in which these notions come in degrees. As a result, strictly speaking, I am not obliged to answer Bar-On's question. However, what drives the question seems to be a resistance to the idea that sufficiently reliable signals could, by themselves, serve to express psychological states when they don't make those states perceptible. And here we do well to reflect on the example that is evidently being offered. If one of these creatures (perhaps a character from the film *The Invention of Lying*) says to one of the others 'I think that's an ugly tie', then he has reported on his belief that it is an ugly tie and, by virtue of making an assertion, expressed the belief that he believes it is an ugly tie. If the compulsiveness of the speaker is supposed to undermine his ability to express himself, some argument for this conclusion will be needed.

¹⁷ See Ekman (1992) for a defence of this approach and a reply to criticisms. For philosophical elucidation and defence of this approach to emotions see Griffiths (1998). Also, many proponents of the natural kinds view of emotions will posit display rules to account for the variation that we observe across cultures in patterns of emotional expression.

turn, supports an approach to emotions as being perceptible under certain circumstances. However, recent experimental psychology has seen a number of serious challenges to the natural kinds view. As discussed in Russell and Fernández-Dols (1997), and detailed in Barrett (2006), the enormous body of data typically appealed to in support of the natural kinds view of emotion underdetermines that view, and furthermore is consistent with other hypotheses, two of which we briefly consider here. On one hypothesis emotions are *idiographic*, with the result that for each person, there is a characteristic pattern of emotional responses that need not be mirrored in others. A has a predictable way of experiencing (physiologically, behaviourally and subjectively) her own anger, fear, surprise and so on, and that characteristic pattern may differ from the pattern exemplified in B for those emotions. As Barrett points out, the experimental data on emotion is consistent with this idiographic view, which of course is logically weaker than the natural kinds view because it does not imply that there are pan-cultural patterns characterizing emotional experience and expression. (The idiographic view plays AE to the natural kinds' EA.)

A more radical challenge to the natural kinds view comes with the suggestion that emotions such as anger, fear, surprise and disgust are folk-psychological categories having no explanatory power. This view would suggest an error theory about common-sense emotional categories, and would replace that system of beliefs with a view of *core affect*, the stream of neurophysiological changes an agent undergoes in response to her environmental influences. Those changes might then be mapped onto a two-dimensional circumplex structure *sensu* Russell (1980). On this approach, there is no such thing as anger construed as a psychological state with causal powers. Instead there are various degrees of internal excitation productive of different kinds of behavioural response, some of which share similarities with the folk notion of anger; likewise for the other allegedly basic emotions. Furthermore, just as we still describe people as being splenetic or bilious, emotion words still provide a convenient way of dividing up affective reality, but would have no role to play in a scientific psychology (Russell, this volume, advocates a similar view).

To what extent would these two rivals to the natural kind view undermine the claim that it is possible to perceive emotions? The ideographic view of emotion is consistent with that claim but would predict, for each person, a narrowing of the range of others' emotions that she is able to perceive. The reason is that the level of interpersonal

variation predicted by the ideographic view would make it challenging for most people to learn the characteristic patterns of emotional expression in very many others. Instead, most people would have the requisite expertise only for those within their social circle, together perhaps with a few others who are known to many.¹⁸

By contrast, the core affect approach would imply that no one perceives any emotions since ‘perceive’ is factive. It would be compatible with the claim that people perceive others *as* experiencing certain emotions. It would also accommodate to a limited extent the intuitive idea that we perceive emotions: no one perceives another’s anger (since there is no such thing), but such an event might very well track a heightened core-affective state with certain properties associated with stereotypical facial expressions and behaviours. Likewise for the other *née* basic emotions.

1.4 Speech Acts

A speech act is an act that one performs by speaker-meaning that one is doing so.¹⁹ For many speech acts, we may in this light distinguish between what a speaker means, and how she means it. *What* one speaker-means in asserting P may be P and no more, but if one is asserting it, one means it in such a way as to be liable to certain kinds of criticism: P might be untrue, or asserted on insufficient grounds, or not believed, and in each case one can be called to account. By contrast, what one speaker-means in guessing that P is the same as in the assertion case, but one means it in a different way. For instance, it is no criticism of a guess that P that it was made in the absence of any compelling evidence in support of that proposition. A similar what/how distinction applies to my predictions of my future actions, as well as to my promises.²⁰

¹⁸ Green (2010: 51–52) considers the possibility that each of us might have competence only to perceive the emotions of those about whom we have acquired expertise. For most of us, the number of people about whom we have the requisite expertise is likely well below the “Dunbar number” of 150.

¹⁹ Speech acts are thus distinct from acts of speech. See Green (2014) for an elaboration. Also, we need not assume that all speech acts express psychological states. It is doubtful, for instance, that appointing or excommunicating does so. Such practices fulfill their social role without relying on expressing speakers’ states of mind.

²⁰ I attempt an elucidation of the sorts of commitment that speech acts engender in Green (forthcoming b).

Why do we have these demanding practices of asserting, promising and the like? We have seen that certain of our psychological states do not admit, or at least do not readily admit of perceptual manifestation. I suggest that culture and individual ingenuity help us overcome this limitation by extending our expressive capacities. (In this section and the next I will discuss the way in which culture does its part; in [Section 1.6](#) I turn to individual ingenuity.) Just as we speak of the institutions of representative democracy and of marriage, we may also contemplate speech-act institutions, which are designed by cultural evolution to aid us in achieving socially valuable aims. Many speech acts do this by enabling us to manifest our states of mind. Apology is a type of act aimed at making amends with someone you have wronged, and in order to achieve this, the institution demands that the speaker who is apologizing regret the offending behaviour. When genuine and not feigned, such regret is the pound of flesh that we offer up in order to make amends: it will constrain your future behaviour by, for instance, tending to make you avoid repeating the offending action in relevantly similar circumstances, as well as by tending to make you cringe when you contemplate that action.²¹ Similarly, assertion is a type of act aimed at conveying information, and the institution of assertion is built to enable this by demanding that speakers believe what they assert. (Because of the importance of trustworthy information transfer, the institution of assertion also demands that their producers have some justification for what they assert.²²)

The notion of a signal invoked in [Section 1.1](#) is drawn from the evolutionary biology of communication. Researchers in that field are concerned with the question under what conditions signalling systems are stable – robust against mutant invaders who signal a state of affairs without its obtaining, and gain an advantage thereby.²³ In so doing these invaders will gain a survival and reproductive edge over their peers that may result in their swamping the population. Once that occurs, the signal will lose its communicative value. We now know that many communication systems are robust against such invasions. They

²¹ Certain circumstances require more for an apology than that the speaker be in a state of regret. Among the Yakuza, for instance, to apologize to a superior that you've offended, you need to cut off a finger in his presence. Here the pound-of-flesh requirement is taken fairly literally.

²² Other members of what Green (2013) calls the 'assertive family' – conjecture, presumption and the like – make related but different demands on their producers.

²³ A central text is Maynard Smith and Harper (2003). For an application of the general theory to the case of bright coloration in tree frogs see Maan and Cummings (2012).

achieve this robustness by enabling signallers to provide not just an indication but strong evidence that the condition in question obtains. This is the source of interest in recent theoretical biology in the phenomenon of 'honest signalling'.

I conjecture that sincerity-demanding speech acts in human societies mirror the stability-supporting features of animal communication systems by enabling us to protect against cheaters. In communities in which speakers are held accountable for what they say, they lose their ability to participate fully in communication if they do not live up to the standards that their speech acts demand. Empty promises, rash predictions and hollow apologies can catch up with us, tarnishing our reputations and compromising our ability to contribute to future conversations and make other valuable linguistic moves. The fact that we put our names on the line when we perform sincerity-mandating speech acts is no accidental characteristic of illocutionary institutions. Instead it is integral to their being able to uphold a viable communicative practice.²⁴

An individual's pattern of malfeasance can result in her being pushed to the edges of communicative life. However, large-scale prejudice can marginalize whole groups of would-be communicators on the basis of gender, race, religion and the like. The result is that a speaker may find it difficult to form the intentions necessary to perform the speech acts she would like to. Though not impossible, it is difficult to form the intentions necessary to carry out a speech act if you do not expect others to take you seriously. But by making speech acts less accessible to certain types of speakers, patterns of injustice will also make it more difficult for them to express their states of mind.

1.5 Charged Language

Perhaps more primitively than the case of speech acts, some words and phrases are designed for the expression of attitudes. A puzzling and controversial case falls within the practice of slurring, in which a speaker denigrates a person on the basis of their membership in a specified group such as the Roma, lesbians or persons with disabilities. We conventionalize such behaviours in various ways including through words: a slurring word is a word or phrase whose conventional meaning

²⁴ Green (2009) develops a view of many speech acts as being analogous to 'handicaps' in the sense of that notion used in the evolutionary biology of communication.

is such as to expresses a derogatory attitude (contempt, mild derision etc.) towards a referent on the ground of its genuine or pretended membership in a certain class. (We might also contemplate the possibility of words or phrases – call them ‘valors’ – that valorize or otherwise promote individuals on the basis of their genuine or pretended group membership.) I say ‘genuine or pretended’ to accommodate the fact that slurs are sometimes used to slur people who are not in fact members of the derogated class. As Jeshion (2013) observes, we may also distinguish between slurring words and words used slurringly. ‘Chink’ is a slurring word, but one can utter the innocuous noun ‘Chinese’ with a contemptuous tone of voice or accompanying facial expression in such a way as to slur Chinese people.

Slurs (valors) have both an expressive role as well as one that contributes to the truth conditions of the sentences in which they occur. An entity falls in the extension of ‘kike’ iff it falls into the extension of ‘Jewish’. (That is, slurs-containing indicative sentences have the same truth-conditions as their ‘neutral counterparts’.) However, it is only in using the former word to refer to someone that a speaker expresses, by virtue of the word’s conventional meaning, a reprehensible attitude towards that person on the basis of his being, or his being pretended to be, Jewish. Further, what is expressed by a slur is not well conceptualized as a further element of speaker-meant content, in the manner either of conventional or conversational implicature. On the most common conception of conventional implicature, one who asserts ‘A but B’ speaker means without saying that there is a contrast or tension between A and B; yet one who slurs another does not speaker-mean a further content beyond what it shares with its neutral counterpart. Rather, just as by calling someone ‘Jewish’ contemptuously I may express contempt without speaker-meaning any content beyond the literal meaning of that word, so too I can use a slur without adding anything to the content that I convey beyond what is conveyed by its neutral counterpart.

According to Richard (2008), holding that slurs have the same truth conditions as their neutral counterparts will yield unacceptable consequences. He writes:

Imagine standing next to someone who uses S as a slur... The racist mutters *that building is full of Ss*. Many of us are going to resist allowing that what the racist said was true. After all, if we admit its truth, we must believe that it is true that the building is full of Ss. And if we think that, we think that the building is full of Ss. We think, that is, what and as the racist thinks. (2008: 13)

The error in Richard's reasoning is readily identified. We may agree with him that if we admit the truth of what the racist says, we must believe that the building is full of Ss. That is, we think *what* the racist thinks. It does not follow that in so doing we think *as* the racist does. We may reply to someone who refers to a Chinese person as 'Chinese' contemptuously, by saying that what they say is true, but their way of putting it is noxious; so too, we might reply to Richard's racist: 'Strictly, what you say is accurate, but your way of putting it is thoroughly offensive.'

Slurs, then, have the same semantic content as their neutral counterparts; they also have a conventionally mandated expressive role making no contribution to the content expressed. What is more, other than quoted occurrences, slurs derogate no matter how deeply embedded they occur in a sentence: uttering the sentence 'John suspected that Mary wondered whether the new cook is a nigger', in a speech act is offensive in spite of the fact that the slurring word is buried under two attitudinatives and one interrogative. It is as if slurs are designed to enable us to outsource the job of scowling with contempt every time we use (rather than merely mention) a certain word in a speech act. This approach to slurs is compatible with their occasionally being appropriated in such a way as to shed their conventionally slurring properties: 'queer' and 'gay' are cases of this kind.

One objection that has been raised to expressivist accounts of slurs is that they have a meaning and, thus, a derogatory and offensive power that remains even when their user does not in fact harbour a derogatory attitude towards their referent (Hom, 2008; Camp, 2013). Camp offers as an example:

I'm glad we have so many spics at our school; they always bring the best food to our fund-raising functions. (Camp, 2013: 332).

A speaker uttering this need not be feeling contempt towards people of Spanish descent. This phenomenon of 'expressive autonomy' – a special case of the broader phenomenon of the autonomy of linguistic meaning (see Green 1997) – is unexceptionable, but should only challenge certain forms of expressivism about slurs. On such forms of this view, one using a slurring word is expressing her contempt (etc.) for the slur's target.²⁵ However, expressivism does not mandate such a

²⁵ Potts, discussing a broader category of 'expressives' of which slurs are a subspecies, writes, "we infer from the speaker's use of [an expressive] that he is in a heightened emotional state *right this minute*" (Potts 2007: 171, emphasis in original). We may see

policy; instead we do better to hold that one using a slurring word in a speech act does something expressive of contempt or a similar attitude towards its target. According to the genitive-or-generic account of expression, such a speaker is thereby expressing contempt towards her target regardless of whether she feels that contempt.

Camp also criticizes expressivist accounts of slurs on the ground that they make a mystery of how one could challenge a speaker's use of a slurring word. If all slurs did was to express one's own current psychological state, this would indeed be an objection to the view. However, on a view of slurs as expressive artefacts, we can make sense of how such challenges are possible. In our reply to Richard, we have already discerned the shape of such a challenge. That may be developed with the thought that implicit in the use of a slurring word is the suggestion that it is an appropriate way of referring to members of a specified group. Absent an interlocutor's demurrer that will normalize this derogatory way of referring to members of that group whether or not all interlocutors would initially have chosen so to refer. As a participant in the conversation, now the only way to pull yourself out of the muck is to explain that *this* way of speaking of the Roma (homosexuals, persons with disabilities, etc.) is offensive and unacceptable.

Jeshion also advocates an expressivist view of slurs, but insists that all slurs derogate in the same way and to the same degree (2013: 308). This surprising pair of theses seems to be in support of her general claim that slurs dehumanize their targets. Such a characterization does not, however, apply across the board. If 'to dehumanize' means 'to encourage others to see an individual or group of individuals as not fully human', then it is true that some slurs dehumanize. However, many slurs do nothing so troubling. 'Limey', 'honky', 'yuppie', 'Pom', 'jock' and even 'cheese-eating surrender monkey' don't dehumanize their targets, but cast aspersions on them in a milder, sometimes even jocular way. Talk of dehumanizing here would be hyperbolic, and once we appreciate this we are free to discern the various attitudes that slurs are designed to express.

Finally, how does this approach to the semantics and pragmatics of slurs help to account for their capacity to offend? I take an artefact or a bit of behaviour to be offensive just in case it would cause offense in decent and right-minded people were they to attend to it properly.

Potts as advocating a form of expressivism that is vulnerable to the expressive autonomy objection.

Portrayal of women in a movie as objects of primarily sexual interest is offensive for precisely this reason. Expressing contempt towards someone who behaves egregiously (the maker of the sexist movie, for instance) need not offend decent and right-minded people: we agree that such egregious behaviour merits a sneer or worse. But expressing contempt towards someone on the basis of their race, gender, sexual orientation or physical disability is offensive according to this criterion. Decent and right-minded people are offended by such behaviour because it shows a readiness to undermine a person's or group's social status on inappropriate grounds.²⁶

1.6 Expressive Figures

Aside from expressing a state by showing others that one is in it, or by making that state perceptible, another important part of our expressive life depends on showing others how an emotion or experience behaves or feels. At times we ascribe expressive properties to natural phenomena: we might speak of a melancholy wind or angry marks on the wall.²⁷ Rather than ascribing affective states to wind or walls, we may instead take a cue from Kivy (2002) in explaining the expressive properties of inanimate objects and processes in terms of their relation to human expression. Some sad music has structural properties – for example, tempo and melody – that are isomorphic to some structural properties of sad human behaviour. Generalizing this observation along the lines of Green (2007), we may see expressive, inanimate phenomena as showing, *inter alia*, how an emotion characteristically appears or behaves. Returning to our case of the melancholy wind of Section 1.1, if someone were to ask ‘what does melancholy sound like?’ we could answer her question by presenting her with the sighing, moaning and wailing sounds of the gale. So, too, the marks on the wall are angry because they may be used to show, and thus do show, how angry people characteristically behave. Although I do not take engagement of the

²⁶ Not all slurs are offensive. Developing our example a bit further, suppose the most influential filmmakers persist in portraying women in demeaning ways. Then the word ‘mogul’ might start to be used regularly with a contemptuous tone, and that tone might be conventionalised over time in such a way as to make that word a slur. ‘Mogul’ might now be used to express contempt towards a person on account of their membership in a group of sexist filmmakers. Here ‘mogul’ would be a slur but not an offensive one.

²⁷ The latter example is from Barwell (1986).

imagination to be any part of what makes a phenomenon expressive, it is nevertheless a symptom of the expressive character of a phenomenon that it tends to engage the imagination in distinctive ways. With the gale blowing outside, it is natural to imagine someone shrieking in grief, and it is natural to imagine these gashes on the wall being made in a rage.

I am suggesting, then, that one way in which an inanimate object or phenomenon may be expressive is by showing how an affective state appears or behaves. But something may also be expressive by showing how an affective state feels. In this final section I will focus on how we use figurative language for this purpose.

A long-recognized aspect of emotions is their tendency to make their possessor construe the world in a certain way (Roberts, 1988). My arachnophobia makes me perceive the innocuous little spider in the corner of the shed as looming and life-threatening; when I close my eyes I imagine it pouncing on me and delivering a jugular-severing bite. Your infatuation with Sergei makes you see his hair as a mane of sable and his eyes as emeralds. That is, our emotions, though not all of our moods, are object-involving, and our feelings towards those objects often display themselves in our perceiving them as being a certain way, whether or not they are, and indeed whether or not we think they are (see also Bennett, this volume). (I don't believe the raisin-sized spider could slit my throat, and you don't believe Sergei's eyes are made of anything but cornea, vitreous humour and the like.) Further, we naturally give voice to these perceivings-as with figurative language. When we do this well, we also enable others to appreciate how we feel. In describing your teenager's bedroom as a disaster area, you convey an image that invites me to imagine looking into her bedroom and seeing broken glass, rebar and smouldering wood. That in turn enables me to appreciate how you feel when gazing into the teen's bedroom. In characterizing my colleague Marissa as a whirling dervish of activity I invite you to imagine her moving so fast that it is difficult to see her. This in turn should help you to understand my feeling of being professionally outdone by Marissa. Much the same effect can be achieved by a simile rather than a metaphor, though less dramatically.

Unlike emotions, moods need not have objects. I might feel anxious on account of over-indulgence in caffeine, and your melancholy stems from a glitch in the level of your serotonin uptake. Yet there need be no answer to the question what I am anxious about or what you are feeling melancholy about. However, because they are like emotions in leading us to construe the world in distinctive ways, we may also express

those moods. I give voice to my anxiety by describing the library as a cacophony of slamming books, creaking chairs and crashes of fingers into keyboards; you share your melancholy by saying that walking is like moving through treacle. In each case we show others how we feel by enabling them to imagine the world as thus described.

Emotions and moods, then, guide us to perceive the world as being a certain way (whether or not it is, and whether or not we believe that it is). Further, we naturally use figurative language to convey how we are thereby perceiving the world as being. Such use of figurative language in turn enables others to understand how we are construing things, and thereby how we feel. The recipe for doing so seems to be: just imagine the object (spider, daughter's room, Sergei, high-achieving colleague) as being like *that*, and you'll have some sense of how I feel. The use of figurative language is thus an example of self-expression exploiting the showing-how dimension of showing.²⁸

In our distant hominid past we may have made our states and processes of mind known to one another primarily by making them available to perception. However, the variety and subtlety of those states and processes has dramatically increased since the advent of language, while the challenge of guarding against deceptive expressions of mind has become ever-present. Human institutions such as those governing speech acts are designed in part to broaden the scope of expression while safeguarding its reliability. Charged language enables us succinctly to categorize others, express attitudes towards them in light of those categorizations and thereby help bring our addressees into alignment with our attitudes even when wholly inappropriate. More flexibly and less convention-bound, metaphor enables us to show others what our affect-laden experience is like. In speech acts, slurs and metaphors, then, we see representation and expression working in tandem to create communicative acts whose power and subtlety is often masked by their familiarity.

References

- Bar-On, D. 2004. *Speaking My Mind: Expression and Self-Knowledge*. Oxford: Clarendon.
 2010. 'Expressing as "Showing What's Within": Critical Notice of Mitchell Green's *Self-Expression*', *Philosophical Books* 51: 212–227.

²⁸ I develop this line of thought further in Green (forthcoming a), with application to the phenomenon of empathy.

- Barrett, L. 2006. 'Are Emotions Natural Kinds?', *Perspectives on Psychological Science* 1: 28–58.
- Barwell, I. 1986. 'How Does Art Express Emotion?', *Journal of Aesthetics and Art Criticism* 45: 175–181.
- Camp, E. 2013. 'Slurring Perspectives', *Analytic Philosophy* 54: 330–349.
- Ekman, P. 1992. 'Are There Basic Emotions?' *Psychological Review* 99: 550–553.
- Gallagher, S. 2008. 'Direct Perception in the Intersubjective Context', *Consciousness and Cognition* 17: 535–543.
- Green, M. 1997. 'On the Autonomy of Linguistic Meaning', *Mind* 106: 217–244.
2007. *Self-Expression*. Oxford: Clarendon.
2008. 'Empathy, Expression and What Artworks Have to Teach', in Hagberg (ed.) *Art and Ethical Criticism*. Oxford: Blackwell, pp. 95–122.
2009. 'Speech Acts, the Handicap Principle, and the Expression of Psychological States', *Mind & Language* 24: 139–163.
2010. 'Perceiving Emotions', *Aristotelian Society Supplementary* 84: 45–61.
2011. 'How to Express Yourself: Refinements and Elaborations on the Central Ideas of Self-Expression', *Protosociology*.
2013. 'Assertions', in Sbisà and Turner (eds.) *Handbook of Pragmatics, Vol. II: Pragmatics of Speech Actions*. Berlin: De Gruyter-Mouton, pp. 387–410.
2014. 'Speech Acts', in E. Zalta (ed.) *Stanford Encyclopaedia of Philosophy* (Summer 2015 Edition).
- in press a. 'Imagery, Expression, and Metaphor', *Philosophical Studies*.
- in press b. 'A Refinement and Defence of the Force/Content Distinction', in Harris, Fogal and Moss (eds.) *New Work on Speech Acts*. Oxford: Oxford University Press.
- Griffiths, P. 1998. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.
- Hatfield, G. 2002. 'Perception as Unconscious Inference', in Heyer and Mausfield (eds.) *Perception and the Physical World: Psychological and Philosophical Issues in Perception*. New York: Wiley, pp. 115–143.
- Hom, C. 2008. 'The Semantics of Racial Epithets', *Journal of Philosophy* 105: 416–440.
- Jacob, P. 2011. 'The Direct Perception Model of Empathy: A Critique', *Review of Philosophy and Psychology* 2: 519–540.
- Jeshion, R. 2013. 'Expressivism and the Offensiveness of Slurs', *Philosophical Perspectives* 27: 308–335.
- Kivy, P. 2002. *Introduction to a Philosophy of Music*. Oxford: Clarendon Press.
- Krueger, J. 2012. 'Seeing Mind in Action', *Phenomenology and the Cognitive Sciences* 11: 149–173.
- Maan, M. and Cummings, M. 2012. 'Poison Frog Colors Are Honest Signals of Toxicity, Particularly for Bird Predators', *American Naturalist* 179: E1–E14.
- Maynard Smith, J. and Harper, D. 2003. *Animal Signals*. Oxford: Oxford University Press.
- O'Shea, J. 2012. 'The Theory of Mind and the Aim of Sellars' Original Myth of Jones', *Phenomenology and the Cognitive Sciences* 11: 175–204.
- Potts, C. 2007. 'The Expressive Dimension', *Theoretical Linguistics* 33: 165–198.
- Richard, M. 2008. *When Truth Gives Out*. Oxford: Oxford University Press.

- Roberts, R. 1988. 'What an Emotion Is: A Sketch', *Philosophical Review* 97: 183–209.
- Russell, J. and Fernández-Dols, J. 1997. 'What Does a Facial Expression Mean?', in Russell and Fernández-Dols (eds.) *The Psychology of Facial Expression*. Cambridge: Cambridge University Press, pp. 3–30.
1980. 'A Circumplex Model of Affect', *Journal of Personality and Social Psychology* 39: 1161–1178.
- Smith, B. 2009. 'Speech Sounds and the Direct Meeting of Minds', in Nudds and O'Callaghan (eds.) *Sounds and Perception: New Philosophical Essays*. Oxford: Oxford University Press, pp. 183–210.
- Smith J. 2015. 'The Phenomenology of Face-to-Face Mindreading', *Nous* 90: 274–292.
- Stout, R. 2010. 'Seeing the Anger in Someone's Face', *Aristotelian Society Supplementary* 84: 29–43.
- Topál, J., Gergely, G., Erdőhegyi, A., Csibra, G. and Miklósi, A. 2009. 'Differential Sensitivity to Human Communication in Dogs, Wolves, and Human Infants', *Science* 325: 1269–1272.
- Weiskrantz, L. 2009. *Blindsight: A Case Study Spanning 35 Years, and New Developments*. Oxford: Oxford University Press.

2. Emotions and Their Expressions

James Sias and Dorit Bar-On

2.1 Introduction

Expressions such as ‘I *see* the anger in his face’, ‘I can *hear* the disappointment in her voice’ and ‘[s]he *felt* him trembling with fear’ imply that the emotions of others are things that we can sometimes perceive – that is, things that we can see, hear and feel. However, one might think that, strictly speaking, the perceptual language here is misleading. One might, that is, think that when we say things like this, we do not really mean that the other’s state of mind is perceptible in the same way that ordinary objects in our visual field are perceptible. Rather, to say, ‘I see the anger in his face’, is to offer a kind of shorthand for something like this: I see him make a facial expression that people commonly make when angry, and so *infer* that he is angry. So, ordinary locutions such as ‘I see the anger in his face’ are misguided. They imply that emotions, like ordinary objects, are perceptible, and that behaviours that express our emotions enable non-inferential knowledge of them by others. But in fact there is no literal perception of emotions; at most what is perceived are behaviours that express the emotions, and our knowledge of others’ emotions is always mediated by our knowledge of other things, so that, even in paradigmatic cases where we take ourselves to recognize someone’s emotion in their behaviour *directly*, what we do is infer the presence and character of the emotion *on the basis of* behavioural evidence.

Note, however, that if our ordinary language is misleading in this way, then it is *also* misleading with respect to the relationship between emotions and their expressions. We often describe faces as happy or relieved, voices as anxious, joyful or scared, bodily demeanours as embarrassed or confident and so on. (Indeed, as one author has

observed, there really appear to be no independently characterizable kinds under which, for example, the facial expressions associated with sadness, or joy, fall, but that ‘can be described without reference to the emotions’ – Peacocke [2004]: 66.) Our descriptions of expressive behaviours suggest that we take the emotions to be embodied in the behaviours that express them – the relief is somehow *right there in his face*, the joy is somehow *right there in her voice* and so on. In general, we take expressive behaviour (at least sometimes) to *manifest* subjects’ emotions in a way that makes it immediately available for others to see, hear or feel. When an emotion is expressed, it is *present in* the facial expression, the tone of voice, the bodily demeanour; it’s as though it rises to the surface, somewhat as the bubbles of boiling water rise to its surface. But if the recognition of others’ emotions is always mediated by the perception or knowledge of other things about them, then this way of conceiving the *expression* of emotions would seem to be mistaken as well.

There are various reasons to be drawn to the ordinary understanding of our expression and recognition of emotions. The common-sense conception, if it were correct, would go some way towards discrediting the oft-maligned Cartesian conception of the relation between mind and body, with its attendant and trenchant problem of our knowledge of other minds. Furthermore, the common-sense conception would seem to mesh well with recent findings from social psychology and neurobiology that point to distinct evolutionary socio-biological functions of the emotions and their expressions in humans and other animals (see, e.g., Damasio 1994, 1999; Ekman 1993, 2003). However, our purpose in this chapter is not to offer a direct defence of the common sense *or* the current scientific conception.¹ Instead, we here wish to examine what philosophical view of the *nature* of emotions could vindicate these conceptions. In other words: What view of the nature of emotions would best enable us to take expressions like ‘I can see the relief in his face/sadness in her eyes’, ‘I can hear the joy/trepidation in her voice’ or ‘I felt the anxiety in her body’ entirely at face value? What would sadness, joy, relief, and so on have to *be* in order for it to be possible to perceive such emotions through expressive behaviour, in a way that does not require treating the behaviour as observational *evidence* from which their presence is to be inferred?

¹ For a recent attempt to defend a perceptual account of (what the author describes as) ‘face-to-face mindreading’, see Smith (2015).

We begin (in [Section 2.2](#)) with a critical discussion of a recent account of emotion expression, according to which behaviours that express emotions enable the perception of the expressed emotions in virtue of forming ‘characteristic components’ from which the emotions are appropriately inferred. That discussion will lead us to lay down a certain desideratum: to accommodate the common-sense view, we will suggest, a theory of emotions should allow that emotions can be made perceptible through expressive behaviour in the sense that such behaviour enables *non-inferential perceptual knowledge* of emotions. In [Section 2.3](#), we offer a brief survey of leading philosophical theories of the nature of emotions. In [Section 2.4](#), we argue that some of these theories are less apt to meet this desideratum and sketch a view of emotions that we take to be a more viable candidate that, moreover, can address some of the issues raised in [Section 2.2](#) concerning the perceptibility of emotions through expressive behaviour.

2.2 Expression, Emotion and Perception

We humans have a great variety of ways of expressing our states of mind. At one end of the spectrum, we have the so-called *natural expressions*, such as yelps, grimaces, gasps, tears and laughter, where both the behaviour and its connection to the expressed states are supposed to be inculcated by nature. The so-called *basic emotions* (e.g., fear, anger, happiness, disgust) provide paradigm examples of states of mind that have natural expressions. The behaviours that express these emotions are unacquired, are thought to be pan-cultural and are the ones we share with at least some other primates (see Hwang and Matsumoto, this volume, for evidence in favour of such a view, and Russell and Jack, both this volume, for criticisms of it).² But, of course, many of our emotions go beyond this category, and among the nonverbal behaviours that express them are facial contortions, gestures and bodily demeanours that are mimicked, or acquired, and are more or less conventional (e.g., giving a hug, jumping for joy, stomping one’s feet, tipping one’s hat, sticking out one’s tongue); yet such behaviours often become ‘second nature’. As individuals, we can also develop idiosyncratic ways of expressing our emotions (see Green 2007: 140–143). And, of course, as language speakers, we have a wide array of conventional verbal means of various types – less or more articulate – for expressing emotions

² For discussion and references, see Bar-On (2013).

and sentiments (e.g., ‘Rats!’, ‘Sorry’, ‘Thank you’, ‘What a nasty thing to do’).

It is useful in this connection to separate expressive performances or *acts* from the expressive *vehicles* used in them. Thus, upon seeing a friendly dog, little Johnny’s face may light up; or he may let out an excited gasp, pointing at the dog; or he may emit a distinctive sound (‘Uh!’), or call out: ‘doggy!’ as he reaches to pet the dog; or he may exclaim: ‘Wanna pet the doggy!’ perhaps without reaching. Johnny’s facial expression and his gasp are natural expressions whereas his eager reaching and subsequent utterances are expressive behaviours he voluntarily or perhaps even intentionally engages in to give vent to his desire to pet the dog. Among the utterances, note, are English *sentences*, which have conventional linguistic meaning, and express propositions in (what Sellars calls) *the semantic sense* (Sellars 1969; see also Bar-On 2004: 216f.). Still, these all seem genuine instances of expressive behaviour (see Green, this volume, for an argument that speech acts are expressive). What renders them so has to do with similarities among the performances or acts, which equally serve to give vent to Johnny’s state of mind. These similarities obtain despite significant differences among the *expressive vehicles* used. One can *give expression to a state of mind* (express in *the mental-state sense*) – for example, one’s amusement at a joke – by laughing (where we may assume that laughter does not stand in a semantic representational relation to being amused), as well as by uttering a sentence with a structured meaning, such as ‘This is so funny!’. These are similar expressive performances, yet they use different vehicles of expression.³

But what is it that makes a performance or act one of *expressing*? An idea that takes its inspiration from earlier philosophical work on expression, including remarks by Wittgenstein, as well as Ayer (1952), is that distinctively expressive performances are not merely symptoms, nor even simply signs that reliably indicate states of mind. Instead, engaging in behaviour that expresses one’s state of mind is making it manifest, or *showing* it (to an appropriately placed and suitably endowed

³ Sellars (1969) distinguishes expressing in the semantic sense from expressing in the *causal* and the *action* senses. Bar-On (2004) distinguishes between an *act* of expressing and its *product* and between the *process* and *vehicle* of expressing, and defends a ‘neo-expressivist’ construal, according to which an avowal such as ‘I’m so glad to see you!’ ‘a-expresses’ *the speaker’s joy* at seeing the addressee, using a vehicle that ‘s-expresses’ the self-ascriptive *proposition* that the speaker is glad to see her addressee (see especially Chs. 6–8). For a recent exposition, see Bar-On (2015).

audience), as opposed to merely reporting, or *telling of* it, on the one hand, and as opposed to hiding it, on the other hand.⁴ The showing involved is not that of a mathematical or logical proof or observational evidence; it's not showing *that* something is the case. It's also not like the modelling involved in various kinds of maps. When confronted with an animal baring its teeth in anger, a child squealing in delight, a man raising an eyebrow in surprise or shrugging shoulders in resignation, we may take ourselves to be *directly witnessing* how things are with the expresser. On the expresser's side, the showing behaviour relevant to expressing is behaviour that springs spontaneously from – and exhibits, displays or directly reveals – the expressed state of mind, as opposed to simply providing information or giving evidence about it (the way, e.g., someone taking an aspirin gives us evidence *that* they are in some kind of pain). On the audience's side, the uptake of the behaviour amounts to some kind of direct recognition of the expressed state, as opposed to requiring, say, inferring *that* the expresser is in the state on the basis of various features of the behaviour coupled with contextual information and background knowledge.⁵

At this point, it seems useful to call upon the idea that at least some behaviour that expresses states of mind shows those states in the sense of *enabling the (literal) perception* of the states themselves (by suitably endowed observers). The idea has found support especially in connection with emotions, which, like some sensations, are reliably associated with 'signature' expressions – typical, species-wide overt behaviours – whose perception, by a suitably endowed observer, can suffice for recognizing the presence and character of the emotion itself. Thus, for example, Mitchell Green (this volume and elsewhere) has recently defended the view that 'it is possible literally to perceive the emotions

⁴ This is a contrast that informed, for example, Ayer's discussion when introducing his influential emotivist view of ethical discourse and distinguishing it from the subjectivist view; he says: 'if I say, "Tolerance is a virtue", I should not be making any statement about my own feelings or about anything else. I should simply be evincing my feelings, which is not at all the same thing as saying that I have them' (1952: 109). For recent developments of this view, see, for example, Bar-On (2004) and Green (2007). See also Bar-On (2010) for a critical discussion of Green, on which the present section partly draws.

⁵ A related distinction is drawn by ethologists and biologists when they describe animals' 'affective displays' as 'merely expressive', meaning that they are directly tied to, and directly manifest, animals' affective states. Such displays are contrasted with *intentionally produced* behaviours that are designed to provide an audience with information about the producer or her environment that the audience can then derive, using contextual clues and other knowledge.

of others' (Green 2010: 45). Here 'literal perception' is to be contrasted with merely perceiving *that* someone has a certain emotion, the latter being a correlative of having been (merely) provided with 'strong evidence for the truth of a proposition' (Green 2010: 46; see also Green's contribution to this volume). For Green, the kind of showing relevant to expression is a matter of *enabling knowledge*, including, but not limited to, perceptual knowledge.⁶ (Importantly, though, it turns out that Green thinks of all perceptual knowledge as a matter of drawing inferences based on evidence. See the pages that follow.)

Green supports the view that emotions (*unlike* cognitive states of mind, such as intentions or beliefs) are made literally perceptible through expressive behaviour by considering the expression of the basic emotions – anger, sadness, happiness, fear, surprise and disgust – which, following Griffiths (1998, 2004), he takes to be 'natural kinds' (Green 2010: 50). Behaviours that express such emotions are said to constitute observable *characteristic components* of the relevant states.⁷ This means that we can think of the perception enabled by emotion expressions on the model of 'part-whole perception': just as we can perceive an apple by perceiving its facing surface, or touch a building by touching part of its exterior wall, so we can perceive someone's anger by perceiving a signature expression on her face, inasmuch as the (angry) facial expression is a component or part of the (whole) emotional state of anger (see Green 2010: 56f.).⁸

In a critique of this proposal, Stout objects that 'if the facing surface along with its observable properties can exist independently of the apple then seeing the facing surface by itself does not amount to seeing

⁶ For Green, expressing one's state of mind *can* also be a matter of showing *that* (in the sense of providing strong evidence sufficient for knowledge) one is in the state. Bar-On (2010) argues that this robs Green's account of being able to ground the aforementioned contrast between showing one's state of mind and (merely) telling, or reporting it.

⁷ Relevant here is pioneering work done by psychologist Paul Ekman (1993, 2003) on human facial expressions (directly inspired by Darwin's seminal 1872 work on expression), as well as more recent applications of it to nonhuman primates (see, e.g., Parr and Maestripietri 2003). Note that a characteristic component need not be an *essential* component, nor do we need to suppose that, for example, being angry necessarily requires showing one's anger through behaviour. So the present suggestion can be divorced from logical behaviourism. For discussion, see Bar-On (2010).

⁸ Green extends the idea of characteristic components to *idiosyncratic* expressions, which we set aside here for simplicity. In recent years, it has been suggested that even linguistic interactions among human beings are best characterized in terms of perception, as opposed to propositional-inferential information processing (see, e.g., Gallagher 2008). However, Green demurs; he thinks that the uptake of *linguistic* expressions can at best be thought to be 'perception-like' (see Green 2010).

the apple. It amounts to seeing something that falls short of the circumstance itself, to use McDowell's phrase, whereas what we were after was a way of understanding how we can see right through to the anger in someone's face' (Stout 2010: 32).

This objection seems wrong-headed to us. Consider again teeth-baring and growling in, say, a leopard, which form the facial and vocal signature of (leopard-)anger. Of course, on a given occasion, an individual animal can show the typical visible and audible signs of anger without in fact being angry, for any number of reasons (and not necessarily through trying to deceive). In that case, no one could *see* the animal's anger by perceiving the teeth-baring or *hear* it by hearing the growl; nor could we say that the animal is *showing* its anger (by growling and baring its teeth). Showing and seeing (and hearing or feeling) are *factive*. But that is consistent with the idea that, on those occasions when the animal *is* angry and engaged in the relevant expressive behaviour, an appropriately placed and suitably endowed observer can perceive its anger by perceiving the facial and vocal expression. Compare: We can see a tree by seeing a component of it – say, one of its branches – even though on occasion, if the branch were severed (but then invisibly propped back up), we might be seeing the branch *without* seeing the tree. If the tree *is* there, attached to the branch, as it were, we can see it by seeing its branch. Likewise, if our animal *is* angry, we can see its anger by seeing a characteristic component of it, despite the fact that it's *possible* for the behaviour to come apart from the relevant state (for discussion, see Bar-On 2004: 298f.).

In response to Stout's objection, Green rightly takes issue with 'McDowell's Constraint', namely, that 'if we take what is available to our experience of another's mind as being compatible with that person's not being in the mental state we impute to her, then we cannot know that she is in that state' (Green 2010: 57).⁹ However, by way of addressing the objection, Green makes what we take to be a false move. He proposes that we should think of part-whole perception – in *both* the case of emotion perception *and* the case of perceiving ordinary objects – as essentially an *inferential* matter.¹⁰ What vindicates our ordinary talk

⁹ Joel Smith has pointed out to us that one might accept McDowell's internalist constraint, but still deny that, when, for example, a perceived branch is attached to the tree, 'all that is available to our experience' as perceivers is the branch, rather than the whole tree.

¹⁰ It may be more appropriate for him to say that the *knowledge* afforded by perception is inferential knowledge. But, as we'll see later, Green speaks of (a certain kind of) inference *both* as being the hallmark of perception and as characterizing the kind of

of literally seeing the (whole) apple by seeing its facing surface is that ‘perception of a facing surface *makes appropriate an inference to the presence* of an entire apple’ (Green 2010: 48, emphasis added). Similarly, what vindicates our ordinary talk of seeing the anger in someone’s face is the fact that here, too, there is an ‘appropriate inference’. Expressive behaviour enables observers to have *perceptual knowledge*, despite the possibility of dissimulation (and other illusions), since ‘in perceiving a characteristic component of an emotion, *I am justified* in some instances *in inferring the existence of that emotion*, and that justification is part of what puts me in a position to perceive that emotion rather than just seem to do so’ (Green 2010: 57, emphasis added). The justification in question is not *logical* or deductive; ‘[t]he characteristic components of emotion that are the touchstones of inferences to whole emotions, and which make affect-perception possible, do not logically entail those emotions’ (Green 2010: 58).¹¹ However, Green doesn’t think that this ‘block[s] knowledge of those emotions’ (Green 2010). The idea here seems to be that, insofar as what we perceive when witnessing the relevant expressive behaviour is a *characteristic* component of the emotion, our inference to the presence of the emotion is sufficiently strong (albeit not deductive) to support knowledge.

Green uses the notion of inference in two different ways. On the one hand, he uses it to characterize the *perceptual* character of the knowledge enabled by expressive behaviour (courtesy of the fact that the relevant behaviour shows- α the emotion rather than merely *showing that*

justification the relevant beliefs enjoy (and which renders them genuine *knowledge*). This seems to us problematic. Frank Jackson makes a similar point in his critique of inferential accounts of perception (see Ch. 1 of Jackson 1977). In particular, he focuses upon the distinction between mediate and immediate objects of perception, and how philosophers often characterize the distinction in terms of the former, but not the latter, depending upon an inference of some sort. Reflecting upon an example from Berkeley, in which a subject hears the sound of a coach, Armstrong argues that perception of the sound is immediate, but perception of the coach is mediate, depending upon an inference from past experience (Armstrong 1961: 20–21). But, Jackson insists, while the subject’s *belief* (or *knowledge*) that he hears a coach may be inferential, his perception of the coach no more depends upon inference than his perception of the sound – provided that the sound is, as a matter of fact, that of a coach. Later in the same chapter, Jackson addresses part-whole conceptions of the mediate–immediate distinction (1977: 13–15).

¹¹ It seems inappropriate to speak of the *characteristic components* (the growling, scowling, etc.) as failing to *entail* logically, as they belong to the wrong category. This way of talking may be endemic to Green’s dual use of the notion of inference that we go on to mention.

the subject has the emotion). The analogy he uses here is to the phenomenon of ‘filling in’, where ‘the visual system “infers” the presence of a triangle from the cues available’ (Green 2010: 49; see also this volume Section 2.2). This analogy – and another one he invokes to the case of blindsight (Green 2010: 47f.; note the scare quotes around ‘infer’) – is intended to draw on the *psychologist’s* notion of inference: automatic, non-conscious, subpersonal, non-propositional, niche-appropriate cognitive or representational transition.¹² At the same time, as we just saw, he invokes the notion of proper (though non-logical) inference in arguing (against Stout, who follows McDowell) that *I* can be *justified* in, for example, *concluding* that my friend is angry, based on experiencing her grimace, where the angry behaviour is supposedly *what gives me reason or evidence* for thinking that she is angry. And I can be said to have *knowledge* of her anger (as opposed to *merely* having evidence), despite the fact that the presence of the grimace provides no logical guarantee for the presence of the relevant emotion (i.e., despite the fact that my evidence is not indefeasible).¹³ Here what seems to be at issue is the *epistemologist’s* notion of inference and its potential role in contributing to justification and thus to knowledge.¹⁴

Limitations of space prevent us from examining more closely whether, or how, a single notion of inference can play these two roles. For our present purposes it will suffice to make the following observation. The combination presented by Green’s view – of expressive behaviour enabling *literal perception* of emotions by providing *evidence sufficient for knowledge, from which observers infer subjects’ emotions* – seems to

¹² As Green elaborates, ‘relative to niche *N*, a bird properly infers (in the current sense of “infer”) the presence of a seed from its facing surface if whole seeds tend with sufficient regularity to be connected to their facing surfaces in *N*’ (2010: 49). See also Green (2010: 50), and this volume.

¹³ We should note that Green doesn’t put the point in terms of reasons, but this way of putting it seems very natural, given his discussion (especially 2010).

¹⁴ Smith (2015: 274f.) makes a related distinction between a psychological and epistemological readings of perceptual accounts of ‘mindreading’ and *contrasts* perceptual accounts with inferential accounts. (It is indeed standard for proponents of perceptual accounts to presuppose that, if others’ emotions are literally perceived, then our knowledge of them is non-inferential.) He notes that there are views that take perception to be inferential, but remarks that it is not clear ‘whether the “inferences” proposed by such accounts are inferential in the same sense as are the inferences with which we are familiar from ordinary conscious reasoning’; his own attempt to support a perceptual account is conducted under the assumption ‘that perceptual experience ... does not involve inference’ (in the relevant sense) (2015: 275).

compromise rather than fully vindicate the common-sense view. That view is aptly described by Wittgenstein when he says:

‘We see emotion.’ – As opposed to what? – We do not see facial contortions and make the inference that he is feeling joy, grief, boredom. We describe a face immediately as sad, radiant, bored, even when we are unable to give any other description of the features. – Grief, one would like to say, is personified in the face. (Wittgenstein 1980: §570)

Now, Green appears driven to the aforementioned combination by (1) taking the ‘characteristic component’ view (inspired by certain neuroscientific work on emotions) to be the best way to address apparent difficulties with the common-sense view, including the ones raised by Stout, and (2) taking the ‘part-whole’ model of perception to be the most promising model for emotion perception. But if the price of the combination is giving up crucial aspects of the common-sense view, as well as some of the key philosophical advantages of invoking it in understanding our knowledge of other minds, it behoves us to look for a better way of understanding the relations between an emotion, its behavioural expression by a subject and the knowledge such behaviour affords in the subject’s observers.¹⁵

Before going on, however, we’d like to mention another puzzle raised by Stout, independently of his dispute with Green regarding the epistemology of perception. Stout thinks that the following two claims, both of which are relatively common and intuitive, together

¹⁵ In the present volume, Green clarifies that the inference made in perception is merely ‘low-level’ inference – this is the *psychological* notion we mention in what follows. However, it seems that appeal to that notion will not be apt in replying to the Stout/McDowell objection. Green notes that ‘inferential processes, of either the high-or-low-level kind, are subject to norms: where there is inference there must be the possibility of fallacy or an analogue thereof’ (47). But, arguably, epistemological norms are different from psychological norms; and from the fact that there are norms governing psychological transitions among subpersonal cognitive states it surely doesn’t follow that such transitions are subject to norms of reason – that it even makes sense to speak of one state providing *good evidence* or *reason* for another. Yet the latter norms are what is relevant to the question whether perception enables knowledge (and if so, of what kind); it is thus the latter that are relevant to Green’s response to the Stout/McDowell objection. Green may be caught in the following dilemma. He can insist that undergoing a ‘low-level’ psychological process of inference *ipso facto* (and contrary to the common-sense view perceptualists are trying to capture) involves one in a ‘high-brow’ inference, whereby our perceiving, for example, someone’s smile always constitutes only the *epistemic basis, evidence, or reason* for believing she is feeling joy. Or he can deny that ‘low-level’ inference necessarily brings ‘high-level’ inference in its train, but then that the account of perceptual knowledge he appeals to in response to the Stout/McDowell objection is not adequate as it stands.

generate a puzzle for anyone who thinks that emotions can be literally perceived:

- (1) Emotions are '*present in*' their expressions – for example, I can see his anger in his face.¹⁶
- (2) Emotions *cause* their expressions – for example, his anger caused him to make that face.

The puzzle is essentially this: How can (1) be true given (2)? That is, how can emotions be *present in* facial expressions and other behaviours (for us to perceive) if they are also the *causes of* those facial expressions and behaviours? If we think of causes and effects as distinct and independent entities, then it is hard to see how (1) could be the case – again, given (2). But if emotions are *not* present in the relevant bodily expressions, then it is incoherent to think that we can *perceive* them *in* those bodily expressions, as per the common-sense idea (Russell, this volume, levels a similar criticism against the Basic Emotion Theory). Stout then proposes an alternative to Green's view that he thinks can solve this puzzle, which adopts an Aristotelian model of cause and effect, according to which causes can be *present in* their effects in virtue of the latter being *processes* in which the former are realized. On Stout's alternative conception, 'emotionally expressive behaviour must be taken to be a process of the emotion, considered as some sort of potentiality, being realised' (Stout 2012: 135).

As we'll see later, we are not convinced that a move to an Aristotelian metaphysics of emotions is forced on us. But in the present context it's worth noting that Stout's 'dynamic' view of emotions (as processes rather than states) does not seem any better suited than Green's to vindicate the idea of literal perception of emotions – *by Stout's own lights*. Recall that Stout's first objection to Green's view was on epistemological grounds: characteristic components can exist without the wholes of which they are parts, so perceiving the former cannot amount to perceiving the latter. One obvious way this problem can arise for emotions is that one can clearly engage in, for example, fear-behaviour, without feeling afraid (on stage, e.g., but not only). However, it's very unclear how taking emotions to be processes rather than states can help address this problem, since it seems that processes, too, can be dissimulated (for discussion of expressive failures, see Bar-On 2004: 226–284). Still, we

¹⁶ Stout says: 'I start with the assumption that emotional states are perceivable features of emotionally expressive behaviour. The key thing about this assumption is that emotional states are taken to be *present in* this expressive behaviour' (2012: 135, emphasis added).

think that Stout is right to draw our attention to the fact that accommodating the common-sense idea about the perceptibility of emotions requires tackling directly the following question:

How should we think of the *nature* of emotional states and their relation to behaviours that express them so that it could make sense to think of emotions (as does common sense) as made literally perceptible and non-inferentially known through expressive behaviour?

2.3 Philosophical Theories of Emotion: A Brief Overview

Philosophical theories of emotion can, for the most part, be divided into roughly three camps: theories in the *feeling* tradition, theories in the *cognitivist* tradition and *hybrid* theories that attempt to combine the virtues of theories in the two main traditions while avoiding their various pitfalls.¹⁷ According to a standard feeling theory of emotion – like the one famously defended by William James – an emotion is a feeling of changes in, and to, the body.

Our natural way of thinking about... emotions is that the mental perception of some fact excites the mental affection called the emotion, and that this latter state of mind gives rise to the bodily expression. My theory, on the contrary, is that *the bodily changes follow directly the perception of the exciting fact, and that our feeling of the same changes as they occur is the emotion*. (James 1950: 449–450)

There is something to like about this account of emotion, but plenty more to dislike. The thing to like is James' rejection of the 'natural way' of thinking about emotions as simply the *causes* of their bodily expressions. Rather than thinking of emotions as merely the mental phenomena that are causally antecedent to the relevant bodily changes – for example, we are first afraid, and then the fear causes our hearts to race and so on – James insists that we think of the bodily changes as somehow involved in the emotion *itself*.

As we said, though, there is nonetheless plenty to dislike about James' account of emotion. Perhaps its most significant flaw is its apparent failure to adequately account for the relationship between emotions and their objects – that is, the intentionality of emotion. Philosophers of emotion are nowadays agreed that emotions have at least two kinds of

¹⁷ In what follows, we set aside work by those philosophers who argue that the class of states commonly referred to as 'emotions' is too heterogeneous to be subsumable under a single theory (see, e.g., Griffiths 1998; Price 2012, this volume; Rorty 1980, 2004).

objects: formal objects and particular objects, or targets.¹⁸ Suppose I ask a friend to accompany me to a job interview, for moral support. He is notoriously a late person, but assures me that he will meet me at a local coffeehouse at a particular time before the interview. Unfortunately, but predictably, he is late again. As I sit there in the coffeehouse, eyes darting back and forth from wristwatch to parking lot, I begin to seethe with anger. In this case, the particular object of the emotion is my friend – *he* is the target of my anger. But as an instance of anger, the formal object of my emotion is the same as any other instance of anger – namely, an *offence*. Every emotion-type has a formal object; in fact, formal objects are what allow us to distinguish one type of emotion from another. Anger and fear are distinguished not so much in terms of how they feel, but rather in virtue of the fact that the former is related to *offence* in the same way that the latter is related to *danger*.

If James is right, and emotions are simply feelings of changes *inside* the body, then it is unclear how they could be directed at things *outside* of the body, like my absent and tardy friend. James himself thought the relationship between an emotion and its object to be a causal one, at one point saying that an emotion is ‘nothing but the feeling of the reflex bodily effects of what we call its “object”’ (James 1884: 194, emphasis added). And more than a century later, contemporary philosophers of emotion in the Jamesian tradition still account for emotional intentionality in causal terms (see, e.g., Prinz 2004: 62f.). But there are good reasons for thinking that this approach to intentionality is mistaken. For one thing, it seems as if emotions are regularly directed at things that could not have caused them, like, for example, fictional objects and future events. And for another, a single emotion might have multiple causes – for example, my anger at my tardy friend might *also* be the effect of too little sleep and

¹⁸ We borrow the terms ‘formal’ and ‘particular’ from Prinz (2004: 62f.), who gets them from Kenny 1963. Solomon picks up the same distinction from Kenny, and adapts it to his own cognitivist theory of emotion (1980: 272–274). And arguably, something like this formal-particular distinction underlies Nussbaum’s discussion of emotional intentionality in Nussbaum (1997). We say that emotions have ‘at least’ these two kinds of objects because some philosophers of emotion add others. Helm, for instance, adds a third object which he calls an emotion’s *focus*, that is, ‘the background object having import to which the target is related in such a way as to make intelligible the target’s having the evaluative property defined by the formal object’ (2010: 310). If I am angry with my kids for tracking mud into the house, for example, then the focus of my anger is something like *having a clean house*, since this makes intelligible why such an action should count as an *offence*. We think this phenomenon is best accounted for by a hybrid theory of emotion that includes some conative component among the parts of an emotion, and defend such a theory in what follows. But we will not pursue this particular issue any further here.

too much coffee on my part – but it is unclear, if emotions are supposed to be the effects of their objects, why only one of these causes should be the intentional object of the emotion.¹⁹

The cognitivist tradition, as it is now known, arose primarily as a reaction to the feeling theory's apparent failure to account for emotional intentionality. According to standard versions of cognitivism, an emotion is just a special kind of evaluative judgement. To be angry with someone, for instance, is to judge that he or she committed some offence; to fear something is to judge that it is dangerous and so forth.²⁰ Judgements, after all, are intentional in much the same way that emotions seem to be – for example, my judgement that John was wrong to steal my car and my anger at him for doing so both have something to do with an *offence*, and are both about *John in particular*. Furthermore, thinking of emotions as judgements may help to explain some of their other features, often included under the umbrella of 'the rationality of emotion' (de Sousa 1987). For instance, we can *justify* our emotions to others, just as we can justify our judgements, by pointing to reasons (Taylor 1975). In such cases, we might say that a person is rightfully angry, reasonably scared and so on. But it seems we cannot say the same for feelings of bodily changes.

As with James' feeling theory, though, there are a number of very serious problems with simple versions of cognitivism. For one thing, if emotions are really just evaluative judgements, then it is difficult to see how emotions and their corresponding evaluative judgements could ever come apart, so that we could have one without the other. But, in fact, this seems to happen quite often, as when a person is afraid of flying despite sincerely judging that it is a very safe way to travel. Another problem is that our ordinary ways of thinking and talking about our emotions implicate the body in various ways, and this appears to be

¹⁹ Prinz's teleosemantic account of emotional intentionality is meant, in part, to explain why emotions pick out some causes as their objects and not others – roughly, because they were *designed* (by evolution and learning) to be caused by those things. Anger is directed at offences, he thinks, because anger was designed to be 'set off' by offences, and not by a lack of sleep or excess of coffee. Even if this works as an account of the relationship between emotions and their *formal* objects, though, we think Prinz's account of the relationship between emotions and their *particular* objects is nonetheless subject to the worries we raise here.

²⁰ This is how cognitivist theories of emotion typically account for the relation between emotions and their formal objects: the formal object of an emotion is just whatever property is attributed to an object in the relevant judgement. All instances of fear, for instance, involve an attribution of the property *being dangerous* (or the like) to some object – so *danger* is the formal object of fear.

lost on simple versions of cognitivism. It cannot be merely incidental to the nature of emotion that we ordinarily use terms like ‘emotion’ and ‘feeling’ more or less interchangeably.²¹ Simple versions of cognitivism, it seems, do not adequately account for the affectivity of emotion.

In light of the serious problems faced by theories in both of the two major traditions, many philosophers of emotion nowadays defend hybrid views according to which emotions have both cognitive and non-cognitive parts (e.g., feelings). Greenspan describes emotions as involving ‘[a]n affective element... [being] brought into connection with an evaluative thought or cognition’ (Greenspan 2010: 547). Roberts models his own theory of emotion after another affective-cognitive phenomenon he calls ‘feelings of construed condition’, for example, the feeling of being triumphant (Roberts 1988, 2003). These states are feelings, but not *mere* feelings, since they also (essentially) involve a kind of cognition that he calls a *construal* – for example, construing oneself as triumphant. Oakley adds a conative element to his hybrid view, conceiving of emotions as complex states composed of beliefs, feelings and desires (Oakley 1992).

We agree that hybridity is a promising way to account for the features of emotions upon which philosophers have often focused – especially their intentionality and affectivity – and take up our own hybrid view in a later section. Our earlier discussion, however, suggests that there may be an *additional* desideratum on theories of emotions, namely, that they accommodate the common-sense idea that others’ emotions are sometimes made directly available to perception through their expressive behaviours. In the next section, we ask whether a theory in one of the two major traditions, or some suitable hybrid, can meet this desideratum.

2.4 What Can Emotions Be That They Might Be Literally Perceived?

As indicated earlier, one of the main reasons for our interest in the possibility of literally perceiving others’ emotional states is that such a

²¹ Nussbaum attempts to avoid this latter problem by adopting a neo-Stoic view of the nature of judgement formation. Very roughly, the idea is supposed to be that the process of forming a judgement is one in which the subject reconciles her whole self to the way a situation seems to her. Sometimes this can be done coolly and calmly (e.g., when we judge that one person is taller than another), but other times, it involves what Nussbaum calls an ‘upheaval of thought’ (e.g., when we judge that we’ve lost someone dearly beloved) (Nussbaum 2001). Whatever an ‘upheaval’ is, it presumably includes some of the very felt bodily changes that Jamesians identify as central to emotional experiences.

possibility would imply that our knowledge of others' emotions can be non-inferential. Our discussion of the expression of emotions has suggested that taking seriously the common-sense idea that expressive behaviour enables literal perception of emotions requires that we think of the behaviour as showing the relevant emotion *in the sense of making it immediately, that is, non-inferentially, available* to suitably placed observers. This is what would vindicate the idea with which we began, namely, that we sometimes perceive, for example, joy, disappointment, fear *in* people's faces, voices, bodies – and, in general, that we perceive the emotions themselves *through* expressive behaviour. However, the views we briefly surveyed in the previous section do not appear readily to accommodate this idea.

Suppose emotions are, at bottom, just evaluative judgements, as per the cognitivist theory. Then the cognitivist would need to explain how an evaluative judgement can be made directly perceptible through expressive behaviour. Cognitive states like judgements, beliefs and intentions are often contrasted with 'embodied' states like emotions and sensations in terms of the former having a very loose and indirect connection to observable behaviour.²² Granted, sometimes someone's facial expression and other behaviours – for example, crossed arms, shaking head and so on – may *tell us all we need to know* about how she has judged a particular circumstance. And, of course, people's *verbal* behaviour can serve to express their judgements ('I think what John did is offensive'), so understanding what they say in relevant circumstances can suffice for knowing what judgements they are making. But it's not clear in what sense we can say in either of these cases that *the judgement itself* is made directly perceptually available to observers. So, if there are good reasons for thinking that emotions *can* sometimes be directly perceived (again, it is not our aim in this chapter to make a case for this), then this may amount to yet another reason for thinking that cognitivists are mistaken to neglect the body in their accounts of the nature of emotion.

²² The 'characteristic component' view, for example, does not seem readily applicable to the case of evaluative judgements. What would the behaviours be that could be thought to constitute 'parts' of evaluative judgements, to express those judgements, and to be perceptible by others (in the relevant way)? Notably, Green limits the 'characteristic component' view and the 'part-whole' perception to the case of emotions, to the exclusion of cognitive states. Expressions of beliefs, intentions and such can at most show *that* someone believes that *p*, intends to *phi* and so on. For critical discussion, see Bar-On (2010).

Might feeling theorists fare better, then, since they focus a great deal upon the somatic aspects of emotional experience? The answer, we think, is that they do not – and the explanation here may reveal something important not only about the conditions necessary for others' emotions to be literally perceived, but also about the very nature of emotion itself. According to a standard version of a (neo-)Jamesian feeling theory, a subject first becomes aware of some 'exciting fact' – such as, say, the presence of a ferocious animal. This information is processed by the brain, which then sends signals to various parts of the body. Those signals then have the effect of producing a pattern of bodily changes – in the case of the ferocious animal, these changes will presumably include things like widened eyes, shortened breath, increased heart rate, tensed muscles and so forth. It is the subject's *feeling* of these bodily changes – or, it may be better to say, it is his feeling of his whole body as these changes occur – that constitutes the emotion.

So there is a pattern of bodily changes, on the one hand, and then there is the subject's feeling of those changes, on the other; and the feeling theorist identifies the emotion with the latter alone. But at least on a standard understanding of what feelings are, by identifying emotions with feelings, feeling theorists actually put us in no better position than cognitivists to accommodate the idea of the literal perceptibility of the emotions of others. As a witness to your emotional episode, I may be able to perceive some of the *changes* you feel – your widened eyes, tensed muscles and so on – but I cannot perceive *your feeling of those changes*. Indeed, some theorists maintain that feelings are by their nature subjective and private (see, e.g., Damasio 1999). So if there is any sense in which I literally perceive your emotions, it is very unclear that feeling theories can accommodate this any better than cognitivist theories.²³

Feeling theories may appear to have a kind of advantage over cognitivist theories in virtue of their attention to the somatic aspects of emotion; but there is a sense in which both theories locate the states that *constitute* emotions 'inside' subjects – they take them to be covert occurrences that bear only an 'external' relation to behaviours that are observable by others. Along these lines, Peter Goldie once complained

²³ By the same token, if, as many assume, someone else's (being in) pain is perceivable, pain, too, should not be conceived to be a state exhausted by the subject's own perception of some damage to her body. (Thanks to Joel Smith for prompting this remark.)

in an interview that theories of emotion in both traditions identify the state which is emotion with something 'inner':

[I]f you think about virtually all emotion theories that are on the go at the moment, they identify emotion with something inner, something mental or bodily, more or less inside the skin. And it strikes me more and more that I'm not absolutely clear why that should be, why emotions should end at a certain point, namely, the inner. (Goldie 2010: 4)

Probably, the matter of emotion perception was not at the forefront of Goldie's mind when he made these remarks. But – again, depending upon how strong the reasons are for thinking that emotions can sometimes be literally perceived – the perceptibility of emotions may be grist for Goldie's mill. Whatever else we think about the nature of emotions, if we want to allow that they are sometimes made directly perceptually available to others, then it seems we have to conceive of emotions as consisting at least in part of something overt, or 'outer' (to play on Goldie's terms).

Along these lines, consider the approach to the nature of emotion taken by neuroscientist Antonio Damasio. Damasio is often counted as a neo-Jamesian, but there is some reason to wonder whether the label really fits. After all, as Damasio conceives them, emotions are *not* feelings. Recall the distinction between the brain-induced pattern of bodily changes, on the one hand, and a subject's feeling of those changes, on the other. Jamesianism is commonly referred to as the 'feeling theory' of emotion precisely because it identifies emotion with the latter of these two phenomena. But Damasio identifies emotion with the former, that is, the brain-induced pattern of bodily changes. He explains,

I see the *essence* of emotion as the collection of changes in body state that are induced in myriad organs by nerve cell terminals, under the control of a dedicated brain system, which is responding to the content of thoughts relative to a particular entity or event. (Damasio 1994: 139)

So, as Damasio construes them, emotions are not feelings; rather, they are things that we feel. This subtle-yet-important difference between Damasio and others in the Jamesian tradition may point us in the direction of the right answer to the question, 'What would emotions have to *be* in order to be literally perceptible through behaviours that express them?' In fact, immediately following the above statement, Damasio himself adds:

Many of the changes in body state – those in skin color, body posture, and facial expression, for instance – are actually perceptible to an external

observer. (Indeed the etymology of the word nicely suggests an external direction, from the body: *emotion* signifies literally 'movement out.'). Other changes in body state are perceptible only to the owner of the body in which they take place. (Damasio 1994: 139)

Now, we think that Damasio is mistaken to *identify* emotions with (just) bodily changes, since it is unclear how such bodily changes could, by themselves, exhibit some of the other features of emotions upon which philosophers have often focused, noted in the last section – for example, their intentionality, rational-evaluative aspects and so forth. But there is nonetheless an important lesson to be gleaned from the perspectives offered by Goldie and by Damasio: if emotions are ever literally perceived in the expressive behaviours of others, then we had better find some way to include bodily changes (some of which are overt, or 'outer', in Goldie's sense) – as opposed to the *feeling of* bodily changes (which is covert or 'inner') – among the constitutive aspects of an emotion.

It is a curious fact about emotions – one that is sometimes thought to distinguish them from other states – that they are often (though certainly not always) simultaneously both private and public affairs. Emotions, like sensations, seem to have something ineliminably subjective or first-personal about them. At the same time, emotions represent paradigm examples of parts of our mental lives that can be *shown to* and *perceived by* others, *in* and *through* certain expressive behaviours. This inner/outer (or private/public) duality, can be adequately captured, we think, by taking seriously the idea that the embodied aspects of emotions – the bodily changes of which Damasio and Goldie *inter alia* speak – constitute part of their very nature. When these bodily changes occur, it is as if the contours of the subject's emotions were being outlined on a glass surface. Those on the outside can trace the contours more or less clearly, and directly, even if they do not experience the drawing of the contours 'from the inside', as does the subject. Similarly, the bodily changes that occur as part of an emotional state may be experienced in one way by the emoting subject and in another way by the rest of us. Roughly: the subject *feels* what we *see*, *hear* and so forth.

Where does this leave us with respect to our search for a theory of emotions that accommodates their literal perceptibility? Theories in both of the two major philosophical traditions struggle to explain how it is that emotions could ever be literally perceived, and for basically the same reason: on each theory, the state identified as the emotion is one that is too far 'inside the head', as it were, to be available to the

perceptual experience of others. Hybrid theories currently on offer may fare better, though, given that they comprise such a diverse lot, it's a bit difficult to assess their prospects. But our observations so far nonetheless place constraints upon such views: to accommodate the possibility of the literal perceptibility of emotions, hybrid views must include among the constitutive parts of an emotion *the very bodily changes* that a subject feels in an emotional state – and not, or not merely, the *feeling of* those bodily changes, as is typically the case.

We now wish to sketch a kind of hybrid theory of emotion that we think satisfies many of the desiderata laid out in the previous sections – most importantly, for our purposes, that of accommodating emotion perception in a way that vindicates common sense – before explaining how such a theory could help to adjudicate the debate described in [Section 2.2](#) between Green and Stout.

In other work, one of us has defended a hybrid view that bears some resemblance to the views of Roberts and Oakley (Sias 2014, 2015). On this view, emotions are complex states consisting of *concerns*, *construals* and *feelings*. Crucially, the view spells out a particular way in which the various parts or aspects of an emotion are related. Roughly, when a subject experiences an emotion, she construes something (an object, a person or a state of affairs) in a way that either clashes or accords with some prior concern of hers,²⁴ and this meeting of construal with concern gives rise to the sorts of feelings that we typically associate with emotions. A state of fear, for instance, involves the subject construing a thing as dangerous, in a way that clashes with her concern for her well-being (or the well-being of someone or something that matters to her), and which gives rise to such affective phenomena as the feeling of an increased heart rate, shortness of breath and so forth.

²⁴ This idea of construals either 'clashing' or 'according' with one's concerns is similar to Robert Gordon's suggestion that all emotions involve either 'wish-frustration' or 'wish-satisfaction'. As Gordon sees it, all of our emotional responses are rooted in prior attitudes towards, or concerning, the things to which we respond emotionally. He prefers to speak of these prior attitudes in terms of *wishing something to be so*. So, for instance, he writes, 'When *S* is angry about the fact that *p*, *S* believes that *p* and wishes it not to be the case that *p*. Since what *S* believes is the contradictory or at least the contrary of what *S* wishes, we can say that *S*'s wish is *frustrated*' (1987: 31). What he calls *wish-frustration* is analogous to what I (Sias) would call a *clash* between *S*'s concerns and the way in which *S* construes something. Gordon continues: 'The same [as was said of anger] holds for all of the negative emotions we have considered. The positive emotions, by the same token, seem to involve a wish that is *satisfied*' (1987: 31). Analogously, on my view, positive emotions involve a kind of *according* of our construals with our concerns.

In light of our discussion earlier in this section, to accommodate the perceptibility of emotions, the view just described would need to be amended so that its focus is less upon a subject's *feeling of* changes in and to her body, and more upon *the bodily changes themselves*. So, on the amended theory, emotions are complex states consisting of concerns, construals *and bodily changes*, some of which are (sometimes) literally perceptible to suitably equipped observers.

How does this theory accommodate some of the other features of emotion, noted during our critical overview of philosophical theories in [Section 2.3](#)? Emotions are intentional, on this view, in virtue of having construals as one of their component parts. The *particular* object of an emotion will always be the object of the relevant construal, and the *formal* object of an emotion is determined by the way in which the object is construed. When I am afraid of a bear, for instance, I construe the bear as dangerous: the bear is the (particular) object of my construal, and the emotion amounts to an instance of fear partly in virtue of the way in which I construe the bear – namely, *as dangerous*. This view avoids the cognitivist's issues with fear-of-flying, however, since construals – unlike judgements – are not belief-entailing. I can construe a thing one way (as dangerous), despite sincerely believing it to be another (perfectly safe). Furthermore, emotions are rationally and morally evaluable, on this view, because concerns and construals are rationally and morally evaluable. Fear of a bear is (or can be) rational, since bears are rightly construed as dangerous. Fear of Yogi Bear, however, would be irrational, since cartoons are *not* rightly construed as dangerous. And anger is a right, or fitting, or morally appropriate response to injustice since this is something about which we ought to be concerned (for a related but distinct account of emotions, see Helm, this volume).

An emotional state, on the amended view, has as one of its constitutive parts or aspects outwardly observable, expressive bodily changes. We should note two important contrasts between our view of the nature of emotion and the one that Green adopts from Griffiths.²⁵ For one thing, it is unclear whether Green counts expressive bodily changes or

²⁵ Griffiths (1998, 2004) describes the basic emotions, or affect programmes, as biologically adaptive affective responses that are (1) informationally encapsulated, (2) triggered by automatic, largely unconscious information processing, (3) culturally universal and (4) present in some non-human animals. Even if all of this is true, though, it is hardly a *theory of emotion*. (Of course, this would not concern Griffiths, as he does not think that there can be such a thing as a 'theory of emotion'. For a critique of Griffiths' views, see Roberts 2003: 14–36.)

merely *dispositions* to expressive bodily changes as among the component parts of an emotion. Of anger, for instance, he writes, ‘it comprises an interrelated set of phenomena: physiological response, behavioural disposition including disposition to facial-expressive behaviour, as well as a cognitive disposition to make judgments of a certain sort’ (Green 2010: 50). If anger is partly composed of a disposition to make a certain facial expression, and not the facial expression itself, then it seems that an angry face could only be *evidence* on which to base an *inference* that someone is angry. Second, Green, following Griffiths, quite explicitly limits his view only to the so-called basic emotions, leaving it unclear how the rest of our emotions – indeed, the vast majority of them – could ever be made perceptible. Our view involves no such limitation, perhaps in large part because we do not share Green’s initial assumption (also borrowed from Griffiths) that emotions are, or ought to be thought of as, natural kinds.

On the amended view, emotions are by their nature expressible, embodied states of individuals. (This doesn’t mean, however, that one cannot be in an emotional state unless one exhibits the relevant expressive behaviours. For one can modulate and even suppress the overt aspects of one’s emotion upon feeling their imminent occurrence.²⁶) There is a clear sense, then, in which emotional states are not subjective, private mental states as traditionally conceived; they are not states ‘*inside*’ individuals. Instead, they are states (whole) *individuals are in*. When you are angry, it is not that there is some internal state in you with overt symptoms that others can use as the outer signs of something inner; rather, there is a certain state *you* are in. (Compare: when a chair has a broken leg, the chair is in a particular state; but being broken is not a state inside the chair. Similarly for non-emotional states of individuals, such as being ready to leave the house. For a proposal along these lines, see Bar-On 2004: 423f.)

We are now in a position to return to the debate between Green and Stout concerning the perception of emotions, discussed in Section 2.2. Our preferred hybrid theory of emotions seems apt to accommodate Green’s basic claim that we perceive another’s emotion by perceiving one or more of its component parts or aspects. However, on our view, when you see a certain facial expression, or hear the tone of someone’s

²⁶ There is increasing evidence that many non-human animals are capable of modulating and suppressing expressive behaviours under a variety of circumstances and for a variety of reasons. For discussion, see Bar-On (2013).

voice, there is no need for you to *infer* that the subject is in the relevant emotional state. It's not as though you perceive one (outer) bit and make a projection to other (inner) bits, to yield the information that the subject has this or that emotion. When water boils, it is in a particular state. But when we observe the bubbles on its surface, we do not infer on that basis the presence of some 'inner' state of boiling; rather, we literally perceive (a part or aspect of) the state the water is in. Similarly, when we observe an emoting subject's facial expression and other expressive behaviours, we do not infer the presence of some 'inner' state of emotion; rather, we literally perceive the emotional state the subject is in, since to perceive a constitutive part or aspect of it is to perceive *it*. Your perception of the other's emotional state, like your perception of the state the water is in, is in this sense direct or immediate. It is not underwritten by an inference that, in turn, *supports* (or enables) your knowledge of the other's emotion. Your knowledge of the subject's emotion, like your knowledge of the water's boiling, is non-inferential.

(Two comments: First, the mere possibility of the part being present without the whole of which it is part does not impugn the possibility of seeing the whole by seeing a part, nor does it imply that the perception of the whole must involve inference. After all, even in cases (if there are such) when one *can* perceive a whole, there is a possibility of illusion. If one is deceived about O, one cannot be said to see (hear, feel) O; but that is just a consequence of the facticity of seeing (hearing, etc.) that we mentioned earlier; it does not entail that perception must be mediated by inference. Second, it may be that, psychologically speaking, even the literal perception of O by perceiving one of its parts requires, at the subpersonal level (as Green maintains), some kind of 'filling in', or transitioning from one representational state to another. Still, there is no reason to take this to imply that the *individual* perceiving the part must be drawing some inference to the presence of the whole of which it is part, or of its unperceived parts.)

Now, although we have rejected James' feeling theory, the hybrid view we just sketched agrees with James on this point: the relationship between emotions and expressive bodily changes is not purely *causal*, but partially *constitutive*. Emotions have the relevant bodily changes as their parts. (As you'll recall, James identified the claim that bodily changes are related to the emotions purely causally as the 'natural way' of thinking about the relation between emotions and their expressions in the body. And he explicitly rejected that view in favour of the one that *involves* those bodily changes *in* emotions.) This means

that we can avoid the puzzle Stout raises for any view that embraces the causal claim along with the claim that emotions are ‘present in’ their expressions. For, with James, we reject the causal claim. Yet we do not think this requires resorting to Stout’s Aristotelian metaphysics of cause and effect. On our proposal, anger is present in a facial expression in virtue of having the facial expression as one of its component parts or aspects.

2.5 Concluding Remarks

Our discussion suggests that an adequate theory of the emotions that pays heed to common-sense ideas about their perceptibility must attend to considerations about expression, perception and the nature of emotions – somehow all at once. For we saw how too narrow a focus on just one of these ‘moving parts’ can lead us astray. In articulating our amended hybrid view of emotions, we have been guided by psychological, epistemological and metaphysical desiderata inspired by the common-sense idea that emotions can be literally perceived in the bodily behaviours that express them. This idea seems to gain support from the neuroscientific findings concerning the socio-biological functions of the bodily changes associated with emotions. Yet traditional views of the nature of emotions, we saw, do not readily accommodate it. However, as we also saw, at least one recent attempt to spell out the idea (the ‘characteristic component’-cum-‘part-whole’ view) seems to compromise rather than vindicate it. If we are right, we can retain many of the advantages of the traditional views, while incorporating recent insights concerning the possibility of perceiving emotions, and without compromising the common-sense idea.

Inevitably, however, what we have said raises a number of issues that we cannot address here. In explaining how our view avoids difficulties we found with the ‘characteristic component’ proposal, we adverted to talk of embodied states *individuals are in* as opposed to (mental) states *in individuals*. Clearly, more needs to be said about what this distinction amounts to, and about the appropriate metaphysics for embodied states of individuals – a lot more than we can say here. We think the humdrum examples we have invoked – of the state of water when it’s boiling, the state a chair is in when one of its legs is broken – as well as additional examples, such as that of the state a shirt is in when it’s stained, or a room when it’s tidy, or the process a building undergoes when it’s being demolished, or a person when she is (or is

getting) ready for a party, are suggestive. Our ordinary talk seems rife with reference to states and processes that have observable aspects or parts, where such reference cannot be readily construed in terms of providing evidence for what is 'inside' or 'hidden'. It remains to be seen, however, to what extent understanding talk of emotions and their uptake along these lines can be made to fit with a proper metaphysics of mind, as well as with an adequate psychology and epistemology of perception. Our hope is, however, that we have provided motivation for further exploration of these issues.

References

- Armstrong, D. 1961. *Perception and the Physical World*. London: Routledge and Kegan Paul.
- Ayer, A. J. 1952. *Language, Truth, and Logic*. New York: Dover.
- Bar-On, D. 2004. *Speaking My Mind: Expression and Self-Knowledge*. Oxford: Clarendon Press.
2010. 'Expressing as "Showing What's Within": On Mitchell Green's *Self-Expression*', *Philosophical Books* 54: 212–227.
2013. 'Origins of Meaning: Must We "Go Gricean"?', *Mind & Language* 28: 342–375.
2015. 'Expression: Acts, Products, and Meaning', in Gross, Tebben and Williams (eds.) *Meaning without Representation: Essays on Truth, Expression, Normativity, and Naturalism*. Oxford: Oxford University Press, pp. 180–209.
- Damasio, A. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Avon Books.
1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Cambridge, MA: MIT Press.
- Darwin, C. 1872. *The Expression of the Emotions in Man and Animals*. New York: New York Philosophical Library.
- de Sousa, R. 1987. *The Rationality of Emotion*. Cambridge, MA: MIT Press.
- Ekman, P. 1993. 'Facial Expression and Emotion', *American Psychologist* 48: 384–392.
2003. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*, 2nd edn. New York: Owl Books.
- Gallagher, S. 2008. 'Direct Perception in the Intersubjective Context', *Consciousness and Cognition* 17: 535–543.
- Goldie, P. 2010. 'Interview with Peter Goldie', *Praxis* 2: 4–12.
- Gordon, R. 1987. *The Structure of Emotions*. Cambridge: Cambridge University Press.
- Green, M. 2007. *Self-Expression*. Oxford: Oxford University Press.
2010. 'Perceiving Emotion', *Proceedings of the Aristotelian Society Supplemental Volume* 84: 45–61.
- Greenspan, P. 2010. 'Learning Emotions and Ethics', in Goldie (ed.) *The Oxford Handbook of Philosophy of Emotion*. Oxford: Oxford University Press, pp. 539–559.

- Griffiths, P. E. 1998. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.
2004. 'Emotions as Natural Kinds and Normative Kinds', *Philosophy of Science* 71: 901–911.
- Helm, B. 2010. 'Emotions and Motivation: Reconsidering Neo-Jamesian Accounts', in Goldie (ed.) *The Oxford Handbook of Philosophy of Emotion*. Oxford: Oxford University Press, pp. 303–323.
- Jackson, F. 1977. *Perception: A Representative Theory*. Cambridge: Cambridge University Press.
- James, W. 1884. 'What Is an Emotion?' *Mind* 9: 188–205.
1950. *The Principles of Psychology*, Vol. 2. New York: Dover.
- Kenny, A. 1963. *Action, Emotion, and Will*. London: Routledge and Kegan Paul.
- Nussbaum, M. 1997. 'Emotions as Judgments of Value and Importance', in Bilimoria and Mohanty (eds.) *Relativism, Suffering and Beyond: Essays in Memory of Bimal K. Matilal*. New Delhi: Oxford University Press, pp. 231–251.
2001. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.
- Oakley, J. 1992. *Morality and the Emotions*. London: Routledge.
- Parr, L. and Maestripieri, D. 2003. 'Nonvocal Communication in Nonhuman Primates', in Maestripieri (ed.) *Primate Psychology: The Mind and Behavior of Human and Nonhuman Primates*. Chicago: Chicago University Press, pp. 324–358.
- Peacocke, C. 2004. *The Realm of Reason*. Oxford: Oxford University Press.
- Price, C. 2012. 'Doing without Emotions', *Pacific Philosophical Quarterly* 93: 317–337.
- Prinz, J. 2004. *Gut Reactions: A Perceptual Theory of Emotion*. Oxford: Oxford University Press.
- Roberts, R. C. 1988. 'What an Emotion Is: A Sketch', *Philosophical Review* 97: 183–209.
2003. *Emotions: An Essay in Aid of Moral Psychology*. Cambridge: Cambridge University Press.
- Rorty, A. 1980. 'Explaining Emotions', in Rorty (ed.) *Explaining Emotions*. Berkeley: University of California Press, pp. 103–126.
2004. 'Enough Already with "Theories of the Emotions"', in Solomon (ed.) *Thinking about Feeling: Contemporary Philosophers on Emotions*. Oxford: Oxford University Press, pp. 269–278.
- Sellars, W. 1969. 'Language as Thought and as Communication', *Philosophy and Phenomenological Research* 29: 506–527.
- Sias, J. 2014. 'Ethical Intuitionism and the Emotions: Toward an Empirically Adequate Moral Sense Theory', *Journal of Value Inquiry* 48: 533–549.
2015. 'Being Good and Feeling Well', *Res Philosophica* 92: 785–804.
- Smith, J. 2015. 'The Phenomenology of Face-to-Face Mindreading', *Philosophy and Phenomenological Research* 90: 274–293.
- Solomon, R. 1980. 'Emotions and Choice', in Rorty (ed.) *Explaining Emotions*. Berkeley: University of California Press, pp. 251–282.
- Stout, R. 2010. 'Seeing the Anger in Someone's Face', *Proceedings of the Aristotelian Society Supplemental Volume* 84: 29–43.

2012. 'What Someone's Behavior Must Be Like If We Are to Be Aware of Their Emotions in It', *Phenomenology and the Cognitive Sciences* 11: 135–148.
- Taylor, G. 1975. 'Justifying the Emotions', *Mind* 84: 390–402.
- Wittgenstein, L. 1980. *Remarks on the Philosophy of Psychology*. Chicago: University of Chicago Press.

3. Expressive Actions

Christopher Bennett

3.1 Introduction: Hursthouse's Puzzle

In a well-known paper, Rosalind Hursthouse argues that certain common, and not irrational, actions cannot be accommodated by the dominant philosophical model of the rational explanation of action (Hursthouse 1991). Examples of the category Hursthouse has in mind would include rumpling someone's hair (out of affection or tenderness); jumping in joy or excitement; destroying something connected with a particular person in anger; covering one's face (in the dark) from shame or fear; 'puffing oneself up' with pride; and caressing the clothes of a loved one in grief.

Hursthouse's claim is that (1) these are examples of actions, since they are intentional rather than merely involuntary reflexes (the agent is in some way in rational control of the way she acts); (2) such behaviour is not irrational, but (3) they are not done 'for a reason' in the sense that there is something that the agent takes as counting in favour of acting thus.¹ It is not the case that, in jumping for joy as the ball goes in the net, I am doing so because I believe that this will bring about a state of affairs towards which I have some pro-attitude.² On these grounds, she thinks, these actions cannot be accommodated by the dominant model of rational explanation of action, where actions

¹ Drawing on a McDowellian formulation, Hursthouse says that there is no description of such action that will reveal the 'favourable light in which the agent saw what he did' (on the assumption that to explain action through reasons is precisely to reveal the good that the agent saw in the action).

² Cf. Davidson: 'R is a primary reason why an agent performed the action A under the description d only if R consists of a pro-attitude of the agent towards actions with a certain property, and a belief of the agent that A, under the description d, has that property' (Davidson 1980: 5).

are explained by reference to the purpose the agent had in so acting. She doesn't question the validity of that model, and therefore terms these actions 'arational'. But although there is no *rational* explanation, she claims, there *is* a good explanation for this action: it was done *out of joy* (and behaviour like that is within the normal range of behaviour to which joy leads). She argues that the way we explain these actions, which the dominant model cannot accommodate, is by seeing them as expressions of emotion. Explaining action as the expression of emotion, however, is quite a different kind of explanation of action from what she takes to be the standard form of rational explanation in which we explain action by citing some feature that an agent took to count in favour of the action.

In this chapter I argue that we can reject Hursthouse's conclusion that explanation of action as the expression of emotion is *sui generis* if we have a sufficiently broad understanding of how features can count in favour of actions. In what follows, I will not question her assumption that the standard form of rational explanation of action involves citing some feature counting in favour of the action to which the agent was responding in so acting. What is distinctive about expressive acts, I will argue, is not that they are spontaneous products of the 'overflowing of powerful feeling', requiring a form of explanation that is arational, but rather that they are done intentionally but *for their own sake* as a constitutive part of doing justice to one's sense of the gravity (or indeed levity) of a situation. One ruffles a child's hair in order to *give form*, on this view, to one's sense of the affection-worthiness of the child (given, no doubt, one's relation to the child as well as the child's intrinsic features). The action is not arational, but can be assessed for its expressive adequacy. Furthermore, the expressive action, on this view, has a purpose – that of doing justice, or giving adequate external form to one's sense of the situation. But at the same time, the claim that these are actions done for their own sake – in the sense that nothing further is achieved by 'doing justice to' or 'capturing' or 'embodying in action' the gravity of the situation, that these actions are not done as a means to a further end – means that we can diagnose the tendency to see these actions as 'not done for a reason'. If one takes it that all reasons are means-end reasons, then it will indeed look as though there are no reasons for the kind of acts in question. But if there are reasons sometimes simply to *acknowledge* or *honour* the gravity of a situation then we can reconcile expressive action with the standard picture of rational explanation of action (for an alternative view, see Helm, this volume).

Of course, it will remain true that many manifestations of emotion remain immune to rational explanation. As Hursthouse notes, we can talk about various such manifestations, including the following: phenomena like blushing, sweating and so on, which may betray the presence of underlying emotional states and which seem beyond any voluntary or rational control; forms of behaviour like smiling, which often occur involuntarily, but which it is possible to stop or to suppress; and the apparently fully intentional actions that she is most centrally concerned with. How far rational explanation extends down this scale remains for further research to determine: it will depend on a more developed theory of the sensitivity of the human body and its behaviour to rational considerations than I have to offer here. However, many types of apparently instinctive behaviour can in fact be better understood as habitual, learned behaviour, so where the line is to be drawn would require careful investigation. Furthermore, as will become apparent further on in the chapter, I don't think that the account I give here explains all behaviour out of emotion, or covers all the emotions. Rather, a guiding thought in this chapter is that while some emotions and their manifestations have as their purpose the strategic role of preserving the organism and alerting it to threats or opportunities in its environment (which may be the best explanation of fear, for instance), others have the role of picking out significant events from the otherwise fleeting play of consciousness, and making them resonate in the individual's life, thus providing the agent with the sense of the inherent importance of these events (think rather of guilt and shame). The role of some emotions, in other words, is that of marking or acknowledging, rather than protecting and promoting. If this is correct, then we would expect the resonance to go quite deep, to be felt and manifested bodily in ways that the individual cannot easily control. It is this category of the emotionally expressive – associated with resonance, which seems non-purposive, and which for that reason Hursthouse and others have found puzzling – that I am interested in here.

3.2 How Could We See Expressive Actions as Carried Out for a Reason?

Hursthouse's argument has had its critics, not just because it seems to suggest that much of human behaviour lies outside the scope of rational explanation, but because it casts doubt on the assumption that

intentional action is action that can be explained in terms of reasons.³ As Betzler puts it:

[A]ctions that are intentional but not done for a particular reason (and therefore not carried out because they have brought about something good or valuable) are a puzzle. They count as actions in one respect, but fail to do so in another. They are under the agent's control, yet not intelligible from the agent's point of view. (Betzler 2009: 273)

To overcome Hursthouse's argument (1) we could deny that this behaviour is more than reflex; or (2) we could argue that it is simply irrational; or (3) we could deny that there is no rational explanation for it (i.e., we could deny that there is no way of explaining it as done for a reason). Alternatively, (4) one might try a 'divide-and-rule' strategy, arguing that some of Hursthouse's examples (say those dealing with anger, joy and other 'violent' emotions) fall into the reflex category, while those that are more like, say, rumpling hair or caressing a dead one's clothes, are done for the sake of some expected good. I think it is fruitless to try to deny that the behaviour in question is action, since although much of it might be impulsive, it is not involuntary: while there may be some cases of expressive action in which questions may be raised about the extent of our control over our action (Raz 1999: 38–42), this is not the case, for example, with ruffling the child's hair out of affection. Furthermore, it seems plausible to think that much of this behaviour is not simply hard-wired but learned (and hence culturally variable: think of the different ways in which men and women learn to express joy: punching the air and yelling is pretty masculine: cf. Raz 1999: 41). Could we argue that these actions are irrational? Some of them might be, as when a fit of temper makes one harm or destroy the thing one loves. However, there are many of these actions that, other things being equal, are not contrary to what one has reason to do (though they can be contrary to reason in certain circumstances, or when the expression takes a particular form). Hursthouse seems right to say that these actions are not irrational simply by virtue of being in the category she is interested in, that is, the category of expressive actions.

However, there might seem to be an obvious way of explaining how these actions could be supported by reason (Smith 1998: 22). After all, the agent clearly acts on certain desires – to jump in the air; to destroy everything in the room that belongs to that unspeakable; to cover oneself up; to touch the loved one's clothes – and the action comes about

³ See, for example, Smith (1998); Raz (1999); Goldie (2000); Döring (2003); Betzler (2009).

because the agent believes that moving her body in certain ways will bring it about that she jumps in the air, destroys everything that belongs to him and so on. Here we have identified something that the agent saw as counting in favour of the action: that, given her desire to X, moving her body in certain ways would count as X-ing. However, the problem with this response as a solution to the puzzle is that this kind of explanation would not really explain anything: certainly it wouldn't be a good explanation of the sort that we want when, say, in holding someone to account, we explain human behaviour in terms of reasons.⁴ The explanation does not bring an end to our questions; it simply makes us ask, *why* did she want to jump, destroy and so on? Hursthouse's problem stems from the fact that this minimal response provides no good answer to *that*.

Another way in which one might try to show the rationality of these actions is by trying to find some more intelligible human purpose that could be served by expressive actions, something else that these actions are performed in order to bring about. We could do this by saying that one acts thus in order to express the emotion. This could be interpreted in two ways: (1) that one acts in order to relieve or vent the emotion, to achieve a kind of satisfaction or psychological equilibrium; or (2) that one acts in order to make one's mental state known to some other party or parties. However, neither of these ways of rationalising expressive behaviour seems right. Expressive behaviour is engaged in, in some sense, for its own sake. It is not primarily a technique that we employ to restore psychological balance or to communicate with others (though, of course, those things may be foreseen effects of the expressive action).⁵

However, from these failed solutions to Hursthouse's puzzle, we can begin to see room for an alternative solution. For it is not an entirely unfamiliar thought in normative ethics that some actions might indeed be inherently 'fitting' and good for their own sake: obviously being

⁴ Smith himself notes that this explanation would be 'distinctly unsatisfying' and claims that the Humean account underpinning it must – and legitimately can – be supplemented by reference to the emotion (Smith 1998: 22).

⁵ The communicative view can be given a further twist if it is interpreted as the view that, although the communication is not carried out intentionally, the aim of communication nevertheless explains why we act as we do, for instance if we see expressive behaviour as evolved signaling behaviour which informs others of our attitudes (Green 2007). Green's view is a sophisticated version of a view I will consider in the pages that follow; however, it does not help to answer the question Hursthouse raises of how we can rationalise expressive action by citing some consideration that the agent took as counting in favour of so acting, which is what I am considering in this section.

good for their own sake would be what counts in favour of those actions (cf. Cupit 1996; and the 'expressive theory of practical rationality' canvassed in Anderson 1995). Of course, this idea needs further explanation. But if there are such acts, we can then ask to what extent her examples of action out of emotion can be understood as responsive to such reasons, for instance, by seeing them as spontaneous and intuitive instances of acts done for their own sake. I will now start to explore a sense of expressive action that does point to there being such reasons: reasons that support our deliberations about how to give form to our sense of what is important in a situation.

3.3 Expressive Action as 'Doing Justice'

At this point, I will introduce an example the personal nature of which I hope will not serve as a distraction. During the preparations for my grandfather's funeral, the funeral director raised the question whether my brothers and I would carry the coffin, or whether he should get some of his staff to do it. We thought it over and decided we would do it. Since that time, I have wondered what reasons guided our choice and whether we made the right decision. Let me run through some candidates and briefly comment, not just on the likelihood that those were the reasons on which we acted, but rather on their normative weight, and the extent to which they support our decision.

One thing that may have guided our decision is simply a sense that that is the done thing at these events, and that we did not want to be out of line with the done thing. We will come back to this sort of consideration in what follows, but as specified these reasons are not very strong. Without some sense of why this is the done thing, or what is important about doing the done thing, this looks like mere conformity.

One way to flesh out the importance of doing the done thing would be to say something about the importance of tradition and ritual, particularly rituals that have remained as they are for hundreds of years. It might be said that there is some importance in the continuity achieved by carrying on in the way that people in one's culture act in these situations, and have acted, since in doing so one is connecting oneself to a community dispersed in time and space.

Another way to explain the importance of doing the done thing has a more local normative source, namely, in the desires and expectations of those attending the funeral, particularly my mother whose father it was; or in general wider social expectations. There is no doubt that this

was a significant consideration weighing with us – and a good reason. However, although we may have done it in part for our mother's sake, there is still a question of what makes this a good reason. We did it, let us say, in part because we took it that her unspoken wish would have been that we did so. But why did she so wish? Was this wish reasonable? We certainly would have done a number of bizarre things for her – up to a point – on that day; but this wasn't one of them. For her to wish that we should carry the coffin seemed quite natural, not at all bizarre. We shared in her sense that this was the right and appropriate thing to do. Saying that we did it for my mother's sake doesn't explain the basis of this sense.

In addition to these considerations, we might also have done it for two other reasons that have become familiar through the course of our discussion: either to achieve psychological harmony by letting some pent-up inner states express themselves; or in order to communicate some attitude to others.

Taking the second, which seems the more plausible, we could sketch the following sort of social psychological story. The reason we carry the coffin is because we want to communicate something to others and this is the conventional vehicle (the done thing) by which to do so. What is it that we want to communicate? And why is it important to us to communicate it? It might be said that we want to communicate either something about our grandfather, or, to delve more deeply into our psychology, something about our membership of a particular social group (that, to paraphrase Erving Goffman, we act so as to reassure people that we can be counted on to behave as we are expected to: Goffman 1971). Why is it important to us to communicate such a thing? On the view that we communicate something about our grandfather, we might do so with the intention of receiving support and confirmation from others. On the social psychological view, we might develop the story to include an account of how the reason for much social behaviour is the maintenance of relations of trust, and that this requires the constant affirmation and re-affirmation of one's commitment to play one's role in the group (and specifically, in some stories, to play one's role wholeheartedly, even when one's interests might be better served by breaking key social rules).

However, it seems as though one reason that we might have had for deciding to carry the coffin isn't captured by any of these reasons. This is that doing so might be an important way to honour the departed. The interpretations of our action that we have just canvassed above all

seek to identify some purpose we seek to achieve in carrying the coffin, some end to which this action is contingently the most effective means. This seems to leave out the thought that this act is one that appeared to us as intrinsically fitting to the situation. Now when a theorist claims that something is 'intrinsically fitting', the suspicion might be raised that they are simply refusing to engage in any further argument or justification of their position. However, I think justification can be given, but simply not justification that explains how the act was a contingently effective means to a further end. The main motivation, I think, is the desire to *do justice* to the importance of our loss. The interpretations we have looked at see the act as being addressed at, or done for the sake of, someone or something other than our grandfather. Of course, there are many interests to be served in a social form as rich as that of the funeral. But surely one natural thought might be that the staging of the funeral and the various components one decides to include within the ceremony have as at least part of their focus the person who has died, and that the justification of these actions should have something to do with that person. That is what would be meant by seeing these acts as an expression of, or vehicle for, one's grief for the loss of that person. One way of interpreting the decision to carry the coffin would therefore be this: that one weighty reason in favour of carrying it has to do with the suitability of that act as a way of capturing or reflecting something of what our grandfather had meant to us, and that as a consequence deciding not to do it might be something that we would regret.

In other words, one possibility is that we interpret the action as aimed at *capturing* or *reflecting* something about the situation we are in, namely, the loss of someone important to us, and this is not done as a means to a further end but for its own sake. Can we say something about what makes carrying the coffin an appropriate vehicle for grief? I think we can make some headway – and in doing so we are unpacking or articulating the normative connection (of fittingness) between our sense of the gravity of the situation, on the one hand (the loss of an intimate), and the expressive action that is selected as the vehicle of the emotion, on the other (the carrying). Let us briefly point to some of the considerations that might bear on this question:

- One thing one might want to say is that the act of carrying the coffin exhibits a certain kind of tenderness and caring. Indeed, it is pretty much the last chance anyone will have to do anything for the departed person. Although, of course, the person is beyond

registering anything about it, the act is a highly intimate one. I had never carried my grandfather before. I was aware of being in proximity to his body. It seemed important to get him safely to where he was going.

- There's also the fact of this being a certain journey, and also being, in a powerful metaphor, the end of the journey. The whole event revolves around saying farewell to what remains of the dead person: it is the last point at which he exists as the object he was. As the last thing to be done for him it felt in the end important that he should be carried by the family and not just by people who we were paid to run the service.
- There is also something about the importance of taking responsibility – shouldering responsibility – and being able to deal with this physically demanding task at a time of emotional strain. Self-mastery is required for the action, since a faltering step, or a loss of self-control, can lead to disaster. In shouldering the burden one takes on the all-important role of guiding the person to their final destination, and shows one can live up to it. The act of physically carrying someone in this desperate situation is an embodiment of taking responsibility, displaying one's own gravity and strength.⁶

I think that these are examples of the kind of considerations that might commonly go into deliberation on a question like the one I have posed. They are considerations of a type that intelligent moral agents in our culture do sometimes, perhaps often, deploy as elements of deliberation,

⁶ This feature is potentially troubling in its relation to conceptions of masculinity and femininity. It is no accident, of course, that, in the culture from which this example is drawn, only men carry the coffin. This is not because of physical strength. It is rather a matter of social role. It has to do with the aspects of self-control and strength, being able to keep one's emotions in check. And this makes the ritual problematic if one rejects those aspects of masculinity. Actions like that of carrying the coffin derive their meaning, and rely for their power, on a network of other meanings. In a society where conceptions of gender roles are problematic, it is likely to be particularly difficult to find expressive actions that have no connection with those roles. I don't think that this means we should abandon expressive acts altogether. But it is likely that many such acts, and the attitudes associated with them, are deeply embedded in our sense of the appropriateness of forms of expression. We may seek to create new vehicles for expression that are more equitable. The problem with this is that such vehicles require depth and resonance, and it can be hard for us to see newly coined acts as capturing what it is they want to capture: they wear their arbitrariness on their face. Hence it may be important to attempt to adapt the older forms as particular elements come to lack the power they presumably once had. Perhaps the carrying of the coffin would be more powerful if it were done by male and female alike.

and offer as aspects of justification. They are considerations that bear on the question of whether an action is an appropriate vehicle for an emotion. In asking whether the action is the appropriate vehicle for the emotion, we have to look at how things look to the person experiencing that emotion. I take it that an emotion has perceptual aspects, that it presents situations to an agent in a distinctive way, such that an agent attends to (or finds absorbing or compelling) certain features rather than others, that she sees those features under a certain characterisation, and that this characterisation brings the evaluative features picked out together according to a certain internal logic or narrative: we might call all this the intentional object of the emotion (the emotion is directed at the object under a certain characterisation: cf. Green, this volume; de Sousa 1980; Roberts 1988; Rorty 1980, who make a similar claim in discussing metaphor). An emotion therefore brings with it a sense of the way in which the situation matters. What it means, on this account, for something to be an appropriate vehicle for an emotion is for it to in some way capture or reflect or do justice to our sense of what matters about the situation (or, those key things that are made compelling by the grip of the emotion) – for it to be fitting to the intentional content of the emotional state.

A question now arises about what it means to do justice to the intentional object of the emotion. The way I want to phrase this question is: ‘what is it for an action to have expressive power in relation to a certain situation?’ We can make some progress on this question by looking briefly at one contribution to the debate in aesthetics about expressive qualities of artworks. We might think that debates in aesthetics are an obvious place to look for an understanding of the notion of expression of emotion. However, aestheticians have long since learned to distinguish expressions of emotion, on the one hand, from expressive qualities, and, on the other, to point out, in criticism of the Romantic view of art as the expression of emotion, that artworks may possess the latter without consisting in the former (Hospers 1954–1955). We might nevertheless think that expressive qualities and expressions of emotion are in some way linked. However, insofar as this idea is pursued, there is some tendency in the modern debate about expressive qualities to see such qualities as reflective of the ‘natural expressions’ of emotion (Davies 1980; Kivy 1989) – as though these latter were themselves self-explanatory. I think, therefore, that we need to look elsewhere if we want to know what it is for something to be an appropriate vehicle

or 'natural' expression of an emotion in the first place. The view I am interested in is almost the opposite of this tendency: it will seek to explain the fact that some behaviour can intelligibly be the vehicle for the expression of an emotion through an explanation of that behaviour's expressive power – where having 'expressive power' does not just mean that this is the behaviour we have an observable tendency to engage in when in these emotional states.

3.4 Expressive Acts, Expressive Power and Symbolism

'Expression' can mean a number of different things in philosophy. For instance, a sentence in English might be said to express a (timeless) proposition; utterances, even fact-stating ones, might be said to express our attitudes; moral language might be said to express our convictions rather than stating facts; behaviour might be said to express emotion; music might be said to express emotion; a person might be said to be expressing himself through certain actions or creations; art, music and poetry might be said to have expressive power. There is clearly a question about whether all these uses have anything important in common (Green 2007: 21–22). But that is not my question in this chapter.

I would like to concentrate at this point on the last of these senses of 'expression' – the sense of expressive power. I want to start with an intuition that we might have about some artworks: that they are a vehicle for which we might reach because we feel that they do justice to our sense of how things are in a way that more prosaic and literal media cannot. I would like to suggest that these are artworks that have expressive power by virtue of their success in capturing or doing justice to our sense of how things are (in some respect). If there is anything right about this thought about expressive artworks, then we might say that they are created for their own sake in the sense that they seek to represent rather than change the world: they have a world-to-mind rather than mind-to-world direction of fit. I further want to suggest that we might think of expressive actions in the same way: that they are actions that we reach for in order to do justice to the (perhaps in some way extraordinary) significance of a situation, but where there is no further purpose to be attained than reflecting how things are. In the case of both artwork and action, I will suggest, this reflecting or capturing or doing justice can only be carried out through the adequate symbolisation of the target.

One good starting point is a distinction between two meanings of 'expression' that is drawn by Richard Wollheim in *Art and Its Objects*.

In the first place, and perhaps most primitively, we think of a work of art as expressive in the sense in which a gesture or cry would be expressive: that is to say, we conceive it as coming so directly and immediately out of some particular emotional or mental state that it bears unmistakable marks of that state upon it... Alongside this notion is another, which we apply when we think of an object as expressive of a certain condition because, when we are in that condition, it seems to us to match, or correspond with, what we experience inwardly: and perhaps when the condition passes, the object is also good for reminding us of it in some special poignant way, or for reviving it for us. For an object to be expressive in this sense, there is no requirement that it should originate in the condition that it expresses, not indeed is there any stipulation about its genesis: for these purposes it is simply a piece of the environment which we appropriate on account of the way it seems to reiterate something in us. Expression in this sense I shall (following a famous nineteenth-century usage) call 'correspondence'. (Wollheim 1970: 47)

One way to understand the difference between Wollheim's two senses of expression is to distinguish the *symptoms*, such as crying or smiling, of some mental state from those objects (and, I claim, actions) that possess expressive power because they seem to capture or reflect some mental state (or the content thereof). It is the latter in which I am particularly interested, since it raises issues of adequacy, appropriateness and inappropriateness that ground normative assessment of different forms of expression.

What Wollheim has in mind in the rather cryptic reference to a 'famous nineteenth century usage of "correspondence"' (unexplained in his text) is the tradition of Symbolism in art and poetry that we can see as initiated by Baudelaire.⁷ Specifically, Baudelaire's poem 'Correspondences' (Baudelaire 1857), which sets out the idea that there are resonant affinities between very different types of object, has been called 'the preliminary manifesto of the French symbolist movement' (Dorra 1994: 10). In fact, this idea was highly influential in the Romantic and post-Romantic period, and not just amongst those who prized irrationalism. For instance, it is the basis of Mill's distinction between poetry and eloquence:

Poetry and eloquence are both alike the expression or utterance of feeling. But if we may be excused the antithesis, we should say that eloquence is heard, poetry overheard. Eloquence supposes an audience; the peculiarity of poetry seems to us to lie in the poet's utter unconsciousness

⁷ For some discussion of these sources, see Taylor (1989: Ch. 21).

of a listener. Poetry is feeling confessing itself to itself, in moments of solitude, and embodying itself in symbols which are the nearest possible representation of the feeling in the exact shape in which it exists in the poet's mind. Eloquence is feeling pouring itself out to other minds, courting their sympathy, or endeavouring to influence their belief or move them to passion or to action. (Mill 1973: 80)

As M. H. Abrams reminds us in his magisterial work on the changes that Romanticism brought about in thinking about the nature and value of art, and as reflected in T. S. Eliot's thought that, insofar as a poem is an expression, it is because it serves as an 'objective correlative' of the state of mind of the artist (Abrams 1953: 25; Eliot 1997).

I think that we can usefully develop the idea that expressive action is symbolic in something like this sense. The idea here is that, in performing an expressive action, one seeks to create an external manifestation that corresponds to one's inner state (or rather, the intentional content or object of that state, the way in which the state presents the situation to the subject), and that one does so simply in order to reflect, mark or acknowledge one's sense of the, in some way extraordinary, nature of the situation. The act is a symbol of the situation – or rather, it is expressively powerful insofar as it succeeds as a symbol of the situation – and the symbol manages to capture something about the situation that couldn't be captured otherwise. One acts intentionally in creating the symbol, but the creation of the symbol is its own end in two ways. First, the symbol reflects the nature of the situation, and thus moulds itself to the way the world appears to be rather than attempting to mould or re-shape the world (world-to-mind rather than mind-to-world direction of fit); and, second, in symbolic action simply to have marked or acknowledged the situation is regarded as a sufficient goal in its own right. Symbolic action of this type succeeds when it reflects the world and does not need to be productive of further good.

It is something like this link between emotion and its expression that Sabine Döring seems to have in mind in her response to the debate initiated by Hursthouse. As she puts it:

Expressive actions are rational insofar as the agent has to distinguish appropriate expressions from inappropriate ones. This is particularly important in cases where the action symbolises the representational content of the expressed emotion. Emotions can be symbolically expressed because they are representations, and they are often expressed in this way because they include the target's import for the subject. In the symbolic case, the rationality involved in expressive action consists in grasping the

symbolic relations between emotional representations and their appropriate expressions. (Döring 2003: 227)

And Döring continues in a vein conducive to my deployment of Wollheim:

The most sophisticated way of symbolically expressing an emotion's representational content is achieved in art... What is at stake here is... the expression's appropriateness and quality as a symbol of the way the world appears to the agent in experiencing the emotion. As the exemplary case of artistic expression illustrates, the question whether an expression of emotion is rational is a question of mind-to-world fit rather than of world-to-mind direction of fit. (Döring 2003: 228)

A fuller exploration of what it is for action to be symbolic in this sense will have to wait for another occasion. However, we can make some initial observations. First of all, symbolic action has to be capable of *referring* to the situation it is about. It is understood thus by the agent, and also by third parties. Hence, despite the fact that expressive action is not performed with the aim of communicating, the medium of expressive action must have at least something in common with language,⁸ and we might therefore say that it is apt for communication.⁹ Second, while expressive actions tend to follow certain patterns, and often become stereotyped, there is clearly at least some room for creativity and imagination in coining new forms or developing old ones. For instance, towards the end of Wim Wenders's film, *Wings of Desire*, the two main protagonists finally meet in a bar, after a long build-up in which they had gradually become aware of each other's existence – and, somehow, fallen in love – but had not yet met; they turn to one another in a moment of strong attraction and fulfilment; one expects them (stereotypically) to kiss; but the female character instead looks at her lover and raises her wine glass in cupped hands, in a semi-sacred gesture that somehow captures much more of the tenderness of the situation than a kiss would have done. Third, we might think of expressive actions as bearing something like a metaphorical relation to the situation they concern. This idea would capture the language-like nature of expressive action, as well as the creative and open-ended process of devising them. But this pregnant idea needs to be made more precise. For one thing, symbols seem to have a higher degree of fittingness and

⁸ This is a point stressed by Nelson Goodman in his discussion of expression (Goodman 1976: 45–50).

⁹ Thanks to Catharine Abell for this way of putting it.

exclusivity to them than the free play of metaphors – for instance, it is at least understandable that I might have felt I had not properly paid my respects to my grandfather if I had decided not to engage in this particular symbolic action. So there might be something to the Romantic idea that symbols have to achieve a degree of organic unity, such that any alteration in a part would destroy the whole.¹⁰ For another thing, the idea of action ‘corresponding’ to one’s experience of a situation by acting as a metaphor for it doesn’t in itself explain the specificity of the symbols that seem to compel our imagination. In the funeral case, for instance, we seem to find those actions symbolically important where we act as though we could still look after the departed. Appeal to metaphor does not explain why this particular metaphor is important. Why is just *that* the kind of thing we want to say about this situation? (Though we should also be aware that there might be little in the way of generalization to say in answer to this question.)

On the assumption that what I have said so far has succeeded in making the line of thought to be pursued intelligible and at least somewhat attractive, I would like to now turn to how this account might be deployed in addressing the puzzle of expressive acts with which we started. First of all, I would like to point out that the line I have pursued here is not the only way of rationalising behaviour out of emotion – and I will situate it amongst some other strategies. Second, I will argue for my view by returning to some of Hursthouse’s examples and showing how they can be seen as performed for the kind of reasons of expressive power that we have canvassed in this section.

3.5 Explaining Action Out of Emotion

First of all, I admit that, even if the story told here succeeds in its explanation of expressive action as rational action, guided by reasons, we

¹⁰ Cf. A. W. Schlegel: ‘In the fine arts too, as in nature, that greatest of artists, every genuine form is organic, i.e. determined by the work’s content. In short, form is nothing other than a meaningful exterior, the articulate physiognomy of each object, undisturbed by accidental intrusions, and therefore giving faithful testimony to the object’s hidden essence’ (Furst 1980: 94). Two stories reflecting this theme are quoted by Aaron Ridley (Ridley 1995: 49). The first is Mendelssohn’s claim: ‘A piece of music which I love expresses thoughts to me which are not too imprecise to be framed in words, but too precise. So I find that attempts to express such thoughts in words may have some point to them, but they are also unsatisfying.’ And the second concerns Schumann’s response to being asked what the piece of music he had just played expressed: ‘He sat back down at the piano and played the piece again, saying “That!”.’

should acknowledge that this is not the only way in which this action out of emotion might be revealed as rational. For instance, in Jean Hampton's remarkable analyses of resentment, malice and spite, she argues that action from emotion has at least the appearance of straightforward means–end rationality *given the beliefs or perceptions that the person in the grip of the emotion is subject to* (Hampton 1988). The spiteful person, on Hampton's understanding, sees himself as having been subjugated in some respect – lowered in the pecking order – and takes aim at those he sees as better off to bring them down to his own level, thereby to assuage his own sense of inferiority by giving him company, as he sees it, at the bottom. Now Hampton's nuanced analysis is complicated by the fact that she argues, first, that this emotional strategy is in fact self-defeating, because it doesn't in fact help the person who takes himself to be at the bottom to have others at the bottom with him – since he is still at the bottom; and, second, that the way of thinking that generates this strategy is fundamentally flawed because based on a competitive Hobbesian view of basic human worth rather than a non-competitive Kantian one. Nevertheless, we can see Hampton as presenting us with a way of seeing emotional behaviour as instrumentally rational, given goals that make sense to the person thinking about things through the lens of a given emotion.

Alternatively, we might follow Sartre in seeing action out of emotion, not as straightforwardly strategic, but rather as strategic given some 'magical thinking' – so, for instance, the angry person lashes out because she thinks – though only 'magically' – that in doing so she can destroy that wealth of overwhelming demands that she cannot otherwise satisfy (Sartre 1962). One way of interpreting this is to take Peter Goldie's position, and to say that the magical thinking just means that the person acts as she wishes she was able to act in the situation – that the behaviour exhibits a wish or a desire that runs deep, but which cannot be allowed to surface given the shackles of civilisation (Goldie 2000). On this type of view, behaviour out of emotion can be seen as means–end rational once we appreciate that the 'ends' in question are provided by certain suppressed desires that the strains of involvement bring to the surface.

Finally, we might take the view that emotional behaviour is prudential even though its prudence is not what motivates the agent, or need be transparent to the agent at all. This is the view that behaviour out of emotion is adaptive – or at any rate was adaptive at that time in our evolutionary history when the basic hard-wiring of human psychology

was being laid down (Prinz 2004). Taking the case of fear as a model, for instance, it might be said that this emotion is triggered by typical situations that may or may not themselves represent danger, and that the person who is afraid, in responding to the fearful, need not necessarily be aware of the fearful as dangerous; but that it is nevertheless in the agent's interests to have this short-cut mechanism of responding to the fearful, because, in that way, the agent is likely to be better kept away from danger.

These are at least some of the alternatives for explaining behaviour out of emotion as rational. In putting forward my own account of emotional behaviour as seeking to do justice by symbolising the sense of the situation, I need not thereby be claiming that it is incompatible with these alternatives. One possibility is that my account might be a good explanation of some types of behaviour out of emotion, while the alternatives better explain others – for instance, I find Hampton's account of spite pretty compelling; furthermore, attempting to explain fear-behaviour in terms of 'doing justice' to the situation doesn't seem very plausible. What my account attempts to explain is a subset of behaviour out of emotion: that which is 'expressive' in the sense of being apparently not carried out for any further end, and which rather has as its end marking or acknowledging the significance of the situation in which the agent finds themselves. As the funeral example suggests, sometimes the best explanation of our action is that such action represents the fitting expressive vehicle, in this sense, for our charged sense of the situation. If some emotions, such as grief, pride, guilt, shame, joy, affection, admiration and respect, but unlike fear, are akin to 'judgements of value' in the sense that they are tied to an individual's sense of what is important – specifically, that they make some events in the individual's life resonate through her consciousness rather than slipping away in the endless stream of one thing after another – we will distort our understanding of the behaviour to which they lead if we attempt to fit them in to the strategic model of fear.

The funeral case provides an example of an expressive action about which one might deliberate and regarding which one might end up with some sense of whether one had 'got it right' in acting as one did. The example was chosen precisely in order to display the kind of thinking that might go on about the appropriateness of the symbolic, expressive qualities of the action. Hursthouse's view is that expressions are actions but not done with a purpose. My point is that this overlooks

an important way in which behaviour can relate to inner states. This relation is taken up when we talk of our *giving expression* to certain emotions. In giving expression to an emotion we select some form of expression rather than another as being in some way fitting or satisfying. So my response is that expressive acts *are* done with a purpose, and that that purpose is, as with the funeral case, to mark or acknowledge or do justice to some extraordinary turn of events, or some extraordinary feature of our situation. We do that by creating or ‘coining’ an action of expressive power. However, the act of acknowledging or marking is not done for some *further* purpose. And this, I claim, is what gives rise to the impression that expressive acts are a distinctive category – that they cannot be done as a means to a further end.

Nevertheless, it is natural to think that there is a gap between the funeral case and the spontaneous expressions of emotion Hursthouse mentions. In the latter case, there may be no prior deliberation or planning or selection; indeed, the actions may be done in the presence of strong emotion that might make at least some types of clear thought difficult – though as Raz points out, sometimes strong emotion brings with it a high degree of self-control rather than a weakening (Raz 1999: 39). However, I doubt that this is a definitive reason to reject my account. Many actions are habitual and spontaneous but nevertheless reason-guided: many actions are intentional, or exhibit what Searle calls intention in action, without being preceded by any prior formation of an intention (Searle 1983); and intentions in action can be guided by an agent’s sense of what counts in favour of the relevant actions. I don’t have a full-blown account of habitual rational action to unveil in this chapter. But I will now give a slightly more detailed defence of the application of the funeral model to spontaneous expression of emotion.

Recall that the guiding thought, on my account, is that this action is carried out to give form to, and do justice to, the person’s sense of the situation, as it is presented by the emotion they are in the grip of. To make my account plausible, we would have to be able (1) to see this action as a suitable vehicle to do justice to the way one sees one’s situation in a fit of jealous rage, and (2) to see the motivation of the expressive action as simply that of marking the person’s urgent sense of the situation. Let’s take (1) first.

Take gouging out the eyes to start with. We should note first of all that the eyes are clearly not an arbitrary target. It is a picture, not just of any person, but of the rival in particular. The action taken is not simply that of crumpling the picture up, but rather is directed at the eyes. The

eyes are not just the windows of the soul, but also, perhaps, something without which the rival could not have her beauty – so there is an element, not simply of maiming, but of disfiguring. Perhaps the rival is to be disfigured in order to make it more likely that the object of your affections will choose you; or perhaps, that possibility already lost, it is in order to ensure that she is brought down to one's own miserable level. In his reading of this example, Peter Goldie suggests that we are acting out suppressed desires, desires transformed by their suppression on to the next most suitable object (Goldie 2000); but it seems unnecessary to ascribe to someone who engages in this action the genuine desire to do this to their rival. It would be enough if we could explain what makes the gouging of the eyes resonant and powerful for the person in the grip of jealousy as an externalisation of their complex emotional attitude: it would not follow that jealousy involved a genuine but suppressed desire to do it for real.

If we now turn to the case of ruffling the child's hair, we can see to start with that this is an action that one would only take to someone smaller, more junior, someone on whom one looks down – it is precisely the action of affectionately looking down. Placing the hand on the head in that way, caressing it, is a good way of capturing this attitude. It is also a well-meant invasion of personal space – ruffling the hair of an adult and stranger is a moderately serious breach of rules of personal distance and respect. So you only ruffle the hair of someone whose status doesn't preclude that kind of physical closeness. With these brief remarks we can perhaps start to see how the action might be intelligibly related to the complex intentional content of the emotion.

Finally, let us take the examples of jumping for joy and puffing oneself up with pride and contrast them with the slumping behaviour we associated with grief or dejection. There is an important pattern of 'up' and 'down' here that seems to have resonance as a metaphor, rather than a mere regularity in behaviour. 'Up' is the position of power and activity, of readiness, of open possibility; 'down' is its opposite, of resignation, passivity, inactivity. This is the symbolism that we might see as underpinning our sense that such behaviour is not just common, but appropriate. These are the reasons that might be seen as informing and guiding these expressive actions, actions through which we give form to the emotional state rather than merely manifesting it as a symptom.

In giving these examples, I am trying to show that expressive behaviour even in the spontaneous case is susceptible of a 'reading' that might show why it would be appropriate to select it as a fitting vehicle

for that emotion if, as we were in the case of the funeral, one is in the business of deliberately selecting a vehicle for one's emotion.

If we turn to (2) now, and ask whether it is plausible to think, in each of these cases, that the motivation for the action is not to bring about any further envisaged good beyond, as Wollheim says, the production of (in this case) an action in which one can recognize and see externalised the content of one's own attitude – what are we to say? The view I have developed here is plausible if the view of emotions on which it rests is plausible. Some emotions, I have claimed, are tied to our sense of the importance of things, making it the case that what has importance stands out, for its own sake. Marking our deep sense of things brings it about that these fleeting episodes resonate in our consciousness and take on a life more enduring than the mundane. The way the child strikes you at that moment, in its combination of sturdiness and fragility, makes you feel both lucky in the moment and already nostalgic for what will pass; the fact that things are going so exceptionally well gives you a sense of well-being and benevolent power such that all things seem possible; the constant nagging thought of the rival's supremacy and, in crucial respects, superiority. These emotions reflect, extend and animate our sense of what is important in what goes on; the fact that the behaviour that expresses such emotions also has the role of marking that sense and making it resonate in the external is simply part of the same role.

3.6 Conclusion

In this chapter, I have suggested a way to solve the apparent puzzle that Hursthouse sets about 'arational actions'. Hursthouse worries that there are actions that are intentional but not done for any reason, and which therefore have to be explained as having been done 'out of emotion' rather than for a purpose. My claim is that this conclusion relies on too narrow a view of the reasons for which actions can be performed. Many such actions, I claim, can be thought of as habitual and spontaneous versions of more deliberated actions such as that which I provide in my example of the funeral, acts that are carried out in some sense for their own sake, in order to honour or do justice to the gravity of the situation. More broadly, this response to Hursthouse raises questions about the relation between emotions and the action that expresses them. I have argued that the role of some emotions is to provide an individual with a sense of the extraordinary importance of certain events

in their life, making those events stand out from the manifold, and that the behaviour that expresses those emotions can be seen as contributing to that role. I have claimed that behavioural expressions of such emotions require a dimension of expressive adequacy, even expressive power, and I suggest symbolism as a way of thinking about where such expressive power is to be found. The present investigation leaves many questions about this view unanswered. But I hope to have made a start on sketching the basis on which a neglected alternative in the literature on emotion and its expression could be defended.¹¹

References

- Abrams, M. H. 1953. *The Mirror and the Lamp: Romantic Theory and the Critical Tradition*. Oxford: Oxford University Press.
- Anderson, E. 1995. *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press.
- Baudelaire, C. 1857. *Les Fleurs du Mal*. Paris: Gallimard, 1999.
- Betzler, M. 2009. 'Expressive Actions', *Inquiry* 52: 272–292.
- Cupit, G. 1996. *Justice as Fittingness*. Oxford: Clarendon Press.
- Davidson, D. 1980. 'Actions, Reasons and Causes', in *Essays on Actions and Events*. Oxford: Oxford University Press, pp. 3–19.
- Davies, S. 1980. 'The Expression of Emotion in Music', *Mind* 89: 67–86.
- Döring, S. 2003. 'Explaining Action by Emotion', *Philosophical Quarterly* 53: 214–230.
- Dorra, H. 1994. *Symbolist Art Theories: A Critical Anthology*. Berkeley: University of California Press.
- Eliot, T. S. 1997. 'Hamlet and His Problems', in *The Sacred Wood: Essays on Poetry and Criticism*. London: Faber and Faber, pp. 81–87.
- Furst, L. (ed.) 1980. *European Romanticism: Self-Definition*. London: Methuen.
- Goldie, P. 2000. 'Explaining Expressions of Emotion', *Mind* 109: 25–38.
- Goffman, E. 1971. *Relations in Public: Microstudies of the Public Order*. London: Penguin.
- Goodman, N. 1976. *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis: Hackett.

¹¹ This paper has been presented to audiences at Sheffield, Leeds and Tübingen, and to a workshop on emotion and expression organised by Catharine Abell and Joel Smith in Manchester. I would like to thank those who asked questions at these events. A number of people have also discussed the key ideas with me, and some have provided detailed comments on previous drafts. I am very grateful for their input. In particular I am grateful to Sabine Döring, Joel Smith, Catharine Abell, Dorothea Debus, Kim Brownlee, Daniel Schwartz, Angie Smith, Rob Hopkins, Yonatan Shemmer, Natasha McKeever, Carl Fox, George Botterill, Jenny Saul, David Owens, Carolyn Price, Julian Dodd, Steven Davies, Gerald Lang, Ulrike Heuer, Matthew Noah Smith, Andrew McGonigal, Chris Megone, Bahadır Eker and Bill Wringe.

- Green, M. 2007. *Self-Expression*. Oxford: Oxford University Press.
- Hampton, J. 1988. 'Forgiveness, Resentment and Hatred', in Murphy and Hampton, *Forgiveness and Mercy*. Cambridge: Cambridge University Press.
- Hospers, J. 1954–1955. 'The Concept of Artistic Expression', *Proceedings of the Aristotelian Society, New Series* 55: 313–344.
- Hursthouse, R. 1991. 'Arational Actions', *Journal of Philosophy* 88: 57–68.
- Kivy, P. 1989. *Sound Sentiment: An Essay on the Musical Emotions*. Philadelphia: Temple University Press.
- Mill, J. S. 1973. 'What Is Poetry?' in Trilling and Bloom (eds.) *Victorian Prose and Poetry*. New York: Oxford University Press, pp. 75–83.
- Prinz, J. 2004. *Gut Reactions: A Perceptual Theory of Emotions*. Oxford: Oxford University Press.
- Raz, J. 1999. 'Agency, Reason and the Good', in *Engaging Reason: On the Theory of Value and Action*. Oxford: Oxford University Press, pp. 22–45.
- Ridley, A. 1995. 'Musical Sympathies: The Experience of Expressive Music', *Journal of Aesthetics and Art Criticism* 53: 49–57.
- Roberts, R. 1988. 'What an Emotion Is: A Sketch', *Philosophical Review* 97: 183–209.
- Rorty, A. O. 1980. 'Explaining Emotions', in Rorty (ed.) *Explaining Emotions*. Berkeley: University of California Press, pp. 103–126.
- Sartre, J-P. 1962. *Sketch for a Theory of the Emotions*. Translated by P. Mairet. London: Methuen.
- Searle, J. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Smith, M. 1998. 'The Possibility of Philosophy of Action', in Bransen and Cuypers (eds.) *Human Action, Deliberation and Causation*. Dordrecht, The Netherlands: Kluwer, pp. 17–41.
- de Sousa, R. 1980. 'The Rationality of Emotions', in Rorty (ed.) *Explaining Emotions*. Berkeley: University of California Press, pp. 127–152.
- Taylor, C. 1989. *Sources of the Self: The Making of the Modern Identity*. Cambridge: Cambridge University Press.
- Wollheim, R. 1970. *Art and Its Objects*. Harmondsworth, UK: Penguin.

4. Emotional Expression, Commitment and Joint Value

Bennett W. Helm

We human beings are social animals, and our social nature is in part revealed through our emotions and their expression: by expressing an emotion, we communicate to others something about our attitudes towards them or towards our circumstances more broadly. Yet standard accounts of the emotional expressions tend to consider a relatively narrow range of relatively simple cases, ignoring more complex cases such as expressions of the reactive attitudes – emotions like gratitude and resentment, approbation and guilt – that are central to our ordinary interpersonal relationships. Indeed, as I shall argue, thinking about what is required in these more complex cases can shed light on how we ought to understand the simpler cases.

One aspect of our social nature as human beings is that we share not only our things and our actions but also our cares and values. Emotions are, we might expect, important for valuing in general, both for our attunement to value and even for things being valuable to us in the first place (Helm 2001). In the case of *joint values*– those values we share jointly with others as members of some group – we might expect emotions to play an even more central role. For to have a joint value together with others in a group, it is not enough simply that each member of the group in fact has that value; rather, the coincidence of these individual values must be non-accidental in that each values something because *we* do, where (other things being equal) one would in some sense be failing as a member of the group if one did not value it. Consequently, there must be some means by which such joint values are *instituted* among all members of the group, and this is plausibly where emotions come in: in addition to enabling us to attend and respond to values generally, we might expect emotional expressions to play an important role in communicating these values to our fellow members and thereby reinforcing

the shared sense of value – and shared sense of community – joint values seem to require. My aim here is to explore that role and thereby to argue for a distinctive account of emotional expression, which I shall call the *commitment account*.

4.1 Evincing and Expressing Emotions

A fair amount of philosophical work in the past fifty years has examined emotional expression. What is somewhat surprising is that in this work very little has been done to articulate the content of what is expressed. Now it might seem that the answer is obvious: emotional expressions express *the emotion*. Yet this is not a very precise answer, and it seems that different accounts understand this in different ways. What we might call *minimal accounts* understand what gets expressed to be merely *that* one has an emotion of a particular emotion type: that one is angry or sad or joyful, for example. Such minimal accounts tend to rest on more biologically oriented accounts of emotion and may include an emphasis on understanding the emotional expressions to be grounded in their contribution to evolutionary fitness (see, e.g., Davis 1988; Ekman 1993; Gibbard 1990, Ch. 7; Kovach and De Lancey 2005). In this way, we might understand emotional expressions to have what Dretske (1986) calls either *natural* or *functional* meaning. Thus, the dog's lip-curl and snarl, because they are caused by its anger, 'naturally' express that the dog is angry in the same way that smoke 'naturally' means fire; or (somewhat more strongly) insofar as the lip-curl and snarl has the function, by virtue of its contribution to the animal's fitness, of carrying this information, the expression 'functionally' means that the dog is angry, in the same way that the ringing of the doorbell 'functionally' means that someone is at the door. However, in each case the particular intentional content of the anger, including who the dog is angry at, is not part of its expressive content, though we might be able to infer this from the dog's further behaviour, such as its moves to get in between you and its food; similarly, a person's tearful face expresses sadness but further context beyond the expressive behaviour is needed to be able to make out what he is sad about. Minimal accounts tend to understand this further behaviour to be motivated by a separate desire (to protect the food) and so as not being a part of the expression of the emotion itself.

By contrast, *emotional content accounts* understand emotional expressions to express the evaluative content of the particular emotion the subject has now (see, e.g., Benson 1967; Döring 2003; Koch 1983). On

this type of view, we understand emotions to be evaluations of their targets, where each emotion type involves a distinctive type of evaluation: its *formal object*, as we might tendentiously call it (for some worries about the notion of a formal object, see Teroni 2007). Thus, the dog's anger at me (the *target*) involves an evaluation of me as offensive, whereas the person's sadness about the smashed vase involves an evaluation of this event as a loss. Consequently, emotional content accounts understand emotional expressions to express the dog's angry feeling that I am offensive or the person's sad feeling that the smashing of the vase was unfortunate.¹ (Some prominent accounts of emotional expression are ambiguous between minimal and emotional content accounts, including that of Alston [1965] and Goldie [2000]; I shall have more to say about Goldie's account in what follows.)

The difference between minimal accounts and emotional content accounts can be found not merely in their understanding of the content of emotional expressions but also in their understanding of how such expressions are to be explained. For minimal accounts tend to understand all cases of emotional expression to be non-intentional bits of behaviour that are simply caused by the emotion, behaviours such as changes in respiration or pulse rate or certain facial or other gestures, such as a wince or fist pounding. By contrast, emotional content accounts understand emotional expressions to be intentional actions that are rationally motivated by the evaluative content of the emotion, and such accounts consequently (and rightly) tend to distinguish such intentional *expressions* from what we might call *evinings*: the emotionally caused bits of non-intentional behaviour that are the focus of minimal accounts (see, e.g., Benson 1967; Helm 2001; Koch 1983; others accept this distinction but describe both as 'expressive actions': Döring 2003). Indeed, it is because emotional expressions proper are not just caused but rationally motivated by the evaluative content of the emotion that we can understand them to express that content itself.²

In understanding emotional expressions to be intentional actions, I do not mean to imply that they are goal-directed actions. Indeed, this

¹ One might wonder whether animals like dogs are capable of such emotional evaluations. Although I would not claim that non-linguistic animals are capable of more advanced emotions, such as those I discuss in §3, it is important not to over-intellectualize the emotions. For some discussion, see Helm (2001), especially Chapters 2–3.

² In distinguishing between evincings and expressions, I do not mean to imply that this distinction is particularly sharp, inasmuch as there is not a sharp distinction between intentional action and non-intentional behaviour: there is a fuzzy boundary between

is part of the lesson to be drawn from Rosalind Hursthouse's examples of what she calls 'arational actions' (Hursthouse 1991). Examples of such arational actions include a variety of emotional expressions: jumping up and down out of joy, kissing or rumpling the hair of someone out of love and tearing one's hair or clothes out of grief (Hursthouse 1991: 58). As Hursthouse rightly argues, none of these cases can be rationally explained in terms of the canonical belief–desire model of intentional action, which understands the action as an attempt (guided by belief) to achieve a goal (specified by the desire); consequently, she concludes, these actions are 'done without reason' and are, in this sense, *arational* (Hursthouse 1991: 66). Yet this conclusion follows only if we accept that belief–desire model as correct, and this seems too hasty. We should not think that all intentional action is *instrumental*, done for the sake of achieving some end, for all that's needed for an action to be rationally explicable is that it has a *point* the explanation reveals to be worthwhile. Indeed, this is the case in each of these examples, and it is the evaluative content of the emotion that provides the explanation: the point of jumping for joy is that it is a celebration, and in feeling joy one feels its target to be a good worthy of celebration. (Note that this explanation will not work for jumping out of excitement: excitement does not have a comparable evaluative content that can reveal such jumping to be worthwhile, and so in this case jumping is an arational evincing of excitement rather than an expression of it.) Likewise, kissing someone or rumpling his hair or, depending on the circumstances, holding him or even just sitting quietly with him all have a point in that they are all ways of being solicitous, and in feeling occurrent love one feels someone as worthy of solicitude. And tearing one's hair or clothes has a point insofar as these are all ways of mourning, and to feel grief is to feel a loss to be worthy of mourning. In each of these cases, the emotional expressions are not means undertaken to achieve the end of celebration, solicitude or mourning; they *are* the celebration, solicitude and mourning. So much the worse for the belief–desire model (Betzler 2007, 2009; Helm 2001).³

them such that it can be indeterminate whether some behaviour is intentional or not and hence whether that behaviour amounts to an evincing or an expressing of an emotion. I will have a bit more to say about this in the subsequent paragraphs. (Thanks to Joel Smith for raising this worry.)

³ Given this understanding of intentional actions, we can start to see how fuzzy the distinction between evincings and expressions is. I have claimed that jumping up and down can be an expression of one's joy insofar as that jumping is celebrating. What about a tennis player's fist-pump after winning a hard-fought point? Or a chess player's

This understanding of emotional expressions as rational but non-goal-directed intentional actions has important implications for our understanding both of the formal objects of emotions and, therefore, of the content of emotional expressions. To feel joy is to feel something to be not merely a good but a good worthy of celebration. To feel grief is to feel something to be not merely a loss but a loss worthy of mourning. Likewise, to feel fear is not simply to evaluate something as dangerous, for such an evaluation may be a part of one's excitement instead; rather, in fear one feels this danger to be worth avoiding and so on. That is, fully understanding the *formal object* of an emotion – the evaluation of the emotion's *target* that is characteristic of a particular emotion type – in many (but not all) cases requires making explicit how that evaluation provides the point of certain sorts of action. This is, I suspect, what is behind Peter Goldie's otherwise puzzling claim that expressions of emotions are to be understood as motivated by desires that are 'primitively intelligible, in the sense that they cannot be explained in virtue of anything else other than the emotion which they are a part of' (Goldie 2000: 28). To say that it is 'primitively' intelligible is to say that no further explanation is possible, and this is precisely what I have denied: the formal objects of emotions precisely are evaluations that can make sense of the worth of certain actions.⁴

One might object that this understanding of the relationship between the formal objects of emotions and the point of emotional expressions is uninformative in a way that is no improvement on Goldie's appeal to primitive intelligibility. For it might seem that the only reason we have to understand the formal object of fear to be not merely danger but danger *worth avoiding* is that fear seems rationally to motivate avoidance

smile on seeing how to force a favorable trade? If jumping up and down is a matter of celebrating, why not also think the fist-pump and the smile are as well, so that all three amount to expressions rather than evincings of joy? There are no clear answers here, but that lack of clarity in the phenomena does not tell against the distinction, especially as the distinction applies to clear cases not in this nebulous middle ground between intentional and non-intentional behaviour.

⁴ This makes emotions sound much like desires; indeed, I understand them to be species of a common genus, *felt evaluations*. For details, see Helm (2001, 2002). Green (1970: 554) tries to understand how emotions can give rise to behaviour in terms of what he calls their 'formal orientation', which he conceives as separate from the emotion's formal object. However, as Green understands it, the formal orientation of an emotion just is an arational disposition to respond – to avoid the danger, for example – and so it explains the relevant behaviour merely as a cause rather than, as I have argued, by rationalizing it. Consequently, on Green's understanding avoidance behaviour is what I have called an 'evincing' of fear rather than an expression of it.

behaviour; yet this is at least *ad hoc*, if not viciously circular. We want to be able to explain in particular cases *why* someone finds this particular danger to be worth avoiding – and whether finding it to be so is reasonable – but the account so far seems to offer nothing beyond Goldie's appeal to primitive intelligibility. Indeed, doing so seems fundamental to being able to maintain the distinction between emotional evincings and emotional expressions.

In response, we must specify a further object of emotions in addition to their targets and formal objects: their focus. An emotion's *focus* is the background object that the subject cares about – that has *import* to her – in terms of which we can understand why she emotionally evaluates the target in terms of the formal object. Consider fear. We often simply assume that the focus of fear is the subject's own physical safety, and yet this is true of only a narrow range of cases, as when I fear the bear that rears up and snarls at me. Yet I may also fear such things as the rabbits hopping around my garden in early spring, the kids playing stickball in the street and getting turned down for a loan. In these cases, my fear of the rabbits, the kids and getting turned down is made intelligible in virtue of the import my tulips, my car and buying a new house have to me. It is because I care about having tulips this spring (or my car, which is parked in the street) and because the rabbits threaten to eat my tulip sprouts (or the kids threaten to damage my car) that my evaluation of them as dangerous is intelligible; yet it is because of the import my tulips (or my car) have to me that this danger is intelligible as worth avoiding and so as potentially motivating intentional action. Likewise, my joy when my loan finally comes through intelligibly involves an evaluation of this as a good worth celebrating because of the import this house has come to have for me. Hence, the tulips, my car and this house are the focuses of these emotions targeting the rabbits, the kids and my getting (or failing to get) a loan.

In short, the import of the focus of the emotion to the subject (and the relationship between the target and the focus) explains and justifies the subject's emotional evaluation of the target, including why this evaluation involves finding certain intentional actions to have a point, thereby legitimizing the rich specification of their formal objects I have provided. Indeed, this is why I have elsewhere understood an emotion to be a *felt evaluation*: an intentional response to the import its focus has to one and thereby to the import its target has given the connection between the target and that focus (Helm 2001). As felt evaluations, emotions are responses to the way in which the import of their focuses

plays out in particular circumstances, such that we can understand particular objects in these circumstances – the emotions’ targets – to be evaluable in terms of the emotions’ formal objects: to be a danger worth avoiding, a loss worth mourning, a good worth celebrating and so on. This emotional evaluation of the target in light of the import of the focus in turn explains and rationally justifies our emotional expressions. Consequently, *emotional expressions* just are the potentially non-goal-directed intentional actions rationally motivated by these felt evaluations of the target in light of the import of the focus.

Thus far I have begun to articulate a particular version of an emotional content account, which begins to extend our understanding of what is expressed by emotional expressions proper. Whereas typical emotional content accounts understand what is expressed to be the emotion, complete with its evaluation of a particular target in light of its formal object, my claim is that emotional expressions convey as well the import the focus of the emotion has to the subject. One might think that this difference is not terribly important, that it merely reflects the difference between my account of emotions (as essentially involving a focus) and more standard accounts of emotions. Yet, as I shall argue, when it is fleshed out in greater detail, the difference is significant for several reasons. First, it enables us to make sense of certain cases of actions that are directly expressive of import without any need of intervening desires or emotions, as I shall argue next. Second (as I shall argue in [Section 4.2](#)), it enables us to make sense of the intuitive distinction between the expression of an emotion and a report about an emotion. Finally (as I shall argue in [Section 4.3](#)), it enables us to make sense of the problem with which I began: the role emotional expression plays in communicating, reinforcing and even instituting joint values.

Consider actions directly expressive of import. Here we must distinguish between particular emotions and *evaluative attitudes*: ways of caring about something, of finding it to have import, as a kind of rational disposition to respond to it in a variety of ways, including with particular emotions.⁵ Thus, caring, loving, valuing and respecting are

⁵ It is important not to confuse an evaluative attitude with a dispositional emotion. Thus, I might fear my neighbour’s dog (or dogs generally) insofar as I have a disposition to feel fear in a variety of circumstances with this dog as their common *target*; this is a dispositional emotion. By contrast, an evaluative attitude is a disposition to respond with a variety of emotion types, all of which have a common *focus*. As I shall suggest in [Section 4.2](#), projectable, rational patterns of emotions and desires with a common focus constitute one’s caring about that focus. In some cases, as with ‘love’, we do not have

evaluative attitudes, and we can persist in having these attitudes even in the absence of particular occurrent emotions. When I hug my son out of love, I may do so as the expression of a particular emotion, such as a feeling of affection. Yet this need not be the case: I need not feel any particular occurrent emotion or have any particular desire that motivates my expressive action. Rather, my love for my son – the particular import he has for me – directly motivates my hugging him. The same goes for my shaking your hand out of respect. At issue here is what Stephen Darwall (2006) calls *recognition respect*, an evaluation of someone as having a kind of standing and authority insofar as he is a member of a certain group, as opposed to what he calls *appraisal respect*, which is an evaluation of someone as in some respect a *good* member of that group in virtue of her conduct or character. Recognition respect, then, is an evaluative attitude – a particular mode of caring – that is presupposed by appraisal respect, an emotion (for details, see Helm 2015). When I shake your hand out of recognition respect, I may do so as the expression of a particular emotion (e.g., of gratitude), but I need not: my expressive action may be directly motivated by the recognition respect I have for you – by the particular import you have for me – without any intervening emotion or desire. Consequently, these actions express these evaluative attitudes of love and respect rather than some mediating emotion.

One might object that there must be an emotion or desire mediating between the import and the expression. In reply, first, even in straightforward cases of emotional expression the evaluative attitude itself can play an explanatory role: we might explain my pacing back and forth while my son undergoes a difficult medical procedure equally well by saying that I'm worried about him and by saying that I love him. Consequently, my pacing seems to be that which expresses *both* my worry *and* my love, which is precisely my point in extending the emotional content account in the way I have. Moreover, second, to postulate an occurrent emotion or desire in the case of my hugging my son would seem *ad hoc* rather than explanatory: the only reason we have to postulate an occurrent (but presumably subconscious) feeling of affection is that I hugged my son, whereas we seem to have all the explanation we need by pointing to my evaluative attitude of love without any

distinct words for the emotion and for the evaluative attitude; so as to make it clear what I am talking about, I shall deviate from common English in using 'love' to mean the evaluative attitude and 'feeling of affection' for the occurrent emotion.

such occurrent feeling. The same goes for my shaking your hand out of respect.

That some actions are directly expressive of import is important for how we understand emotional expressions. For there is no fine line between actions directly motivated by import and actions motivated by a mediating emotion. In between cases of my hugging my son as a direct expression of my caring for him – of the import he has for me – and cases in which a wave of affection overcomes me so that my hugging is clearly an emotional expression lies a continuum of cases in which I experience fainter to stronger occurrent emotions of affection. Given this continuum, it would seem that there must be a common core to what is expressed, namely, the import he has for me, even if in some of these cases there is an emotion that is expressed as well. In short, there is a close link between the expression of an emotion and the expression of an evaluative attitude. Getting clear on what that link is will help clarify the nature of emotional expression, and that is my aim in [Section 4.2](#).

4.2 Expressions and Reports

I have been arguing that emotional expressions express not merely that one has an emotion of a particular type directed at some target but also the subject's evaluative attitude towards the focus, the import it has to the subject. This is still too vague, in part because we need to distinguish between *expressing* an emotion (and the underlying evaluative attitude) and merely *reporting* that one has it. This is a distinction found in Ayer (1952: Ch. 6), though of course Ayer's point was to argue that unlike reports that one has an emotion, emotional expressions, which include value statements of various kinds, are not literally meaningful and so cannot be either true or false. William Alston (1965: 16–17) criticizes this view, arguing that linguistic reports – assertions that I have a particular emotion – can just as well be expressions of those emotions as 'cries, groans, squeals, writhings, looks, and tones of voice'. Alston is clearly right to call our attention to the expression of emotions in linguistic propositions as no less central cases of emotional expressions than are interjections or non-verbal bits of behaviour: saying 'I'm angry with you' can express one's anger just as well as directing expletives your way. Yet Alston also seems to miss part of Ayer's point in making the distinction. For, although I do not accept his broader positivism, Ayer does seem to be right to say that in expressing an emotion we are

doing something other than merely reporting that we have it. Thus, I might try to explain my recent rude behaviour by saying 'I'm angry with you', immediately following that with an apology: 'but I realize I shouldn't be, and I'm sorry'. Here my statement of my anger is quite different from my expressing my anger by swearing at you or slamming the door in your face because it is *merely* a report of my psychological state that, in this case explicitly, falls short of an endorsement of that state, and this can be true even when my statement is uttered in a way (through gritted teeth, say) that itself evinces my anger. The question is: what more are we doing when we express an emotion than when we merely report having it?

Answering this question requires understanding more clearly how both emotions and emotional expressions are more or less rational responses to the import of our circumstances, such that these responses can come apart from how we think we ought to respond, thereby making room for the possibility that we can merely report without endorsing those responses. Here I turn to my general account of caring or import for two reasons: it is central to understanding the rationality of emotions and emotional expressions in a way that can provide an answer to the question of what we are doing when we express an emotion other than issuing a report; and these together will provide a framework in terms of which we can understand (as I aim to do in [Section 4.3](#)) the role of emotional expressions in making possible our sharing import jointly. For my claim will be that in certain cases in which joint import is at stake, emotional expressions express *our* attitude towards this import, an attitude that involves certain normative claims on our emotional responses towards this import more generally, where it is these normative claims that ensure that the coherence among our emotional responses is non-accidental.

I have spoken interchangeably of a subject's caring about something and its having import to her. As I have argued elsewhere (see, e.g., Helm 2001, especially Chs. 2–3), these are two ways of describing the same phenomenon, albeit from subjective and objective perspectives, respectively. Consider first the subjective side: caring. To care about something is to have a particular evaluative attitude towards it – to find it to be worthy in various ways of one's attention and action. This means that one must be disposed to attend to it and act on its behalf, but not merely this, for caring is an attitude of *commitment* to attend to and act on behalf of its object, such that one thereby *ought* to attend and act on its behalf. Consequently, one ought to feel a broad range of

emotions and desires with that object as their focus: insofar as you care about your tulips, you ought to worry that the cold, wet spring will rot the bulbs, fear that the rabbits will eat the young shoots, want to protect and fertilize them, be relieved when the application of coyote urine succeeds in keeping the rabbits away and feel joy and satisfaction when the blooms finally emerge. Moreover, this pattern of emotions and desires is itself rationally structured (given one's fear, one ought to feel the subsequent relief; given one's worry, one ought to feel the subsequent joy etc.) in such a way that we can understand this overall rational pattern of emotions and desires to be one's caring commitment itself. Hence, my claim has been that to *care* about something just is for it to be the focus of such a projectable, rational pattern of emotions and desires.

We can understand this same phenomenon objectively, from the perspective of import. That something has import to one makes rational certain kinds of response, including emotional, desiderative and behavioural responses: something that has import to one is something to which one rationally ought to be emotionally, desideratively and behaviourally responsive. Indeed, that the focus of an emotion has import to one is a condition of the warrant of that emotion: my fear that the rabbits will eat the clover is unwarranted if the clover does not have import to me. In this way, the import of the focus is that to which one responds in having these emotions; put another way, emotions just are intentional feelings of import, a mode of attunement to import *impressing* itself on one in feeling.⁶ Here we can begin to see that this talk of the way in which something's having import makes rational certain sorts of responses is another way of describing the normative commitments one undertakes in caring about it: such commitments just are commitments to having the sort of emotional, desiderative and behavioural responses that, from the objective perspective, we can understand to be made rational by its import. This is, roughly, why I have argued that something's having import to a subject just *is* that subject's caring about it.

This understanding of caring and import as subjective and objective perspectives on the same phenomenon has an important implication for emotional expression. I said earlier that emotional expressions express the import the focus and the target have to the subject; this is a

⁶ Indeed, I have argued, this is crucial to understanding the phenomenology of emotions; see Helm (2002).

description from an objective perspective. We can equivalently describe what is expressed by emotional expressions from the subjective perspective by saying that it is the subject's evaluative commitment to the focus and thereby to responding appropriately to that which bears on this focus positively or negatively – to the emotion's target. We might sum these two perspectives together in a single statement in saying that emotional expressions express a commitment to the import of their focus and, in the circumstances, of their target.

How does this conclusion bear on the question of whether in expressing an emotion we are doing something other than making a report? My brief argument at the beginning of this section that emotional expressions cannot be reports hinged on the idea that we need to distinguish between merely reporting that one has a particular emotion and endorsing the evaluation implicit in that emotion, claiming that emotional expressions fall more on the side of endorsement. I have now begun to clarify this in terms of the idea that emotional expressions are commitments to import. Yet it is important not to over-intellectualize such expressions. In general, we might think that to express a commitment is not merely to report that one has it; it is rather to undertake or reaffirm or endorse that commitment. Thus, when I express my belief in global climate change – my commitment to its truth – I am undertaking or endorsing that commitment, an endorsement that itself has normative implications (see, e.g., Brandom 1994). Nonetheless, it should be clear that emotional expressions are not the *undertaking* of commitments to import, for as responses to import they rationally presuppose that one has such a commitment. Moreover, emotional expressions are not exactly *endorsements* of such a commitment, for the idea of endorsement, tied up as it is with the idea of explicit judgement and choice, seems to imply that this represents the subject's own point of view, whereas this need not be true of emotional expressions: I can find myself unable to prevent the expression of an emotion I nonetheless wholeheartedly repudiate, as in cases of phobias. So what room is there to say that emotional expressions are something more than reports?

The answer is that unlike these other cases of the expression of commitment and unlike reports, emotional expressions of a commitment to import are not normally independent of that import itself: normally there is a constitutive relationship between emotional expressions and the import they express, a constitutive relationship that is lacking in mere reports. (I say 'normally' to allow for the sort of phobia cases just mentioned; however, such cases must be exceptional,

lest they undermine the rational pattern of responsiveness to import in terms of which something's having import to one is intelligible.) For emotional expressions are a central part of the rational pattern of attention and action that make something intelligible as worthy of one's attention and action and so as having import to one at all. Thus, other things being equal one would not genuinely care about some aim unless one normally celebrated successes (and anticipated their promise) and mourned failures in its pursuit – unless one expressed one's joys, hopes, sorrows and so on. In this sense, we might say, emotional expressions are the *enacting* of import, of our evaluative commitment. Consequently, emotional expressions are something more than mere reports precisely because, other things being equal, such enactions are rationally required in a way that makes them normally part of what constitutes that import, that evaluative commitment, itself.

This understanding of emotional expressions marks a clear break with emotional content accounts, and to highlight this difference I shall call this a *commitment account* of emotional expression.

4.3 Emotional Expressions and Joint Value

So far I have argued in favour of this commitment account of emotional expression by distinguishing between emotional expressions and reports and by identifying similarities between emotional expressions of import and cases of actions that are themselves directly expressive of import. I now turn to a final argument in terms of the role emotional expression plays in communicating, reinforcing and even instituting joint values.

As I indicated at the outset, in order for values to be shared jointly within a group, it is not enough for members of that group each to have that value, for the coincidence of their values must be non-accidental, such that each values this at least in part because *we* do. The complication is that there are potentially many types of groups that value things jointly, and precisely how those values are shared jointly differs from one type of group to the next. I have discussed two such groups in my earlier work: plural agents (Helm 2008, 2010) and communities of respect (Helm 2015). For brevity I shall discuss only the latter here, though the basic point carries over to other cases: the expression of emotions is a central part of what establishes and reinforces the interpersonal, rational connections among emotions that constitute those

joint values. To understand what I mean by a community of respect, I shall begin with the reactive attitudes.

The notion of the reactive attitudes was introduced by Peter Strawson as the ‘natural human reactions to the good or ill will or indifference’ one person shows to another (1962: 195). Strawson understands there to be three types of reactive attitudes: the *vicarious reactive attitudes*, including indignation and approbation, are the responses of a ‘witness’ to the good or ill will one person (the ‘perpetrator’) shows to another (the ‘victim’);⁷ the *personal reactive attitudes*, including resentment and gratitude, are the responses of the victim to the good or ill will the perpetrator shows to him; and the *self-reactive attitudes*, including guilt and self-approbation, are the responses the perpetrator herself feels in response to the good or ill will she shows to the victim.

Part of the interest of the reactive attitudes as a class of emotions is that they are ways of praising or blaming each other and thus holding each other responsible (or, in the case of self-reactive attitudes, one-self taking responsibility) for what we do. Indeed, Strawson and many who have been influenced by him, myself included (Helm 2015), have argued that we should understand what it is for someone to *be* responsible at least in important parts in terms of the ways she is *held* responsible through the reactive attitudes. Yet for the reactive attitudes to be ways of praising or blaming each other, they must at least in paradigm cases be able to *communicate* this praise or blame to others, including but not limited to the perpetrator. Of course, merely to feel a reactive attitude is not on its own to communicate anything; that reactive attitude must be expressed.

As Stephen Darwall (2006) and Gary Watson (2008) have noted, what is communicated through the expression of reactive attitudes is not merely some evaluation of the perpetrator’s conduct but also the *demand* for a suitable response to that evaluation; as Darwall puts it, such praise or blame comes with an ‘RSVP’ (2006: 40).⁸ Thus, when I harm you, your expression of your resentment (or others’ expression of their indignation) not only communicates your (or their) blaming attitude towards me but also *calls* on me to respond with, for example,

⁷ This language of ‘perpetrator’ and ‘victim’, of course, most naturally fits cases of negative reactive attitudes in response to harm; nonetheless, I shall use it in the positive cases as well, as when someone is the victim of the perpetrator’s generosity.

⁸ For the moment I shall follow Darwall in understanding expressions of reactive attitudes as involving such a demand, though I shall ultimately provide an alternative that I think makes sense of a broader range of cases.

feelings of guilt and attempts at reparation. The same goes for the expressions of self-reactive attitudes: my expression of guilt likewise calls on the victim to respond with at least an acknowledgement of my contrition. As Darwall (2006) rightly notes, what is demanded through the expression of reactive attitudes is in general *recognition respect*: the recognition of our status as having *dignity*, this distinctive kind of import. Moreover, not just anyone can make such a demand of another; rather, to make such a demand one must have the appropriate *authority*, an authority that stems from one's own dignity.

It is commonly assumed that such dignity is one's dignity *as a person*, so that the relevant authority, responsibility and norms to which one is held responsible are all moral in character. This is, I believe, a mistake that blinds us to the ways in which we hold each other responsible in a variety of non-moral contexts. Thus the members of a family might hold each other responsible to certain norms that delimit part of their common way of life and to which they would not hold others – not because these others do not know about the norms (for they might) but because, not being members of the family, they do not have the *standing* as bound by its norms (Helm 2015). Likewise, outsiders do not have the authority to hold family members responsible to the norms of the family: although they can point out that a family member has notably upheld or violated one of the family's norms, this falls short of praise or blame and does not involve any demand for a response. In general, the requisite dignity, standing and authority at issue in the 'call' of the reactive attitudes, and so the recognition respect that is thereby demanded, is the dignity, standing, authority and respect they have as members of such a community – a family, a church, a club, a political group, a nation or even the moral community of all persons. Such communities are *communities of respect*.

The upshot is that the members of a community of respect must jointly value not only the norms by which they regulate their behaviour but also each other as members of the community. These values are linked: in valuing the norms, they do so only insofar as they care about the members as those bound to these norms (without which the caring about the norms itself would have no point), and in caring about the members – in respecting their dignity as members – they do so at least in part by caring about both their standing as responsible to the norms and their authority as able to hold others responsible to these norms. Given this, what can establish and maintain the connections among members such that these values are intelligibly the ones

they hold jointly? My contention is that a central part of the answer involves appeal to the call of the reactive attitudes, a call that depends on the commitment account of emotional expression that I have been developing. To see this, it will be useful to consider the nature of this call from both the objective and subjective perspectives we can take towards import – towards the joint value the members and norms have to each other.

Begin with the objective perspective, understanding the reactive attitudes to be responses to the import of one's circumstances: to the import to us community members of both the relevant norm and of each other as fellow members. For example, assume that within a community of philosophers I interrupt your talk with an *ad hominem* attack, so that you come to resent me and express this resentment after the talk by glaring at me and giving me a cold shoulder. Insofar as your resentment is an intentional feeling not only of the import we have as fellow members but also and thereby of the import of my norm violation in these circumstances, it might seem (mistakenly, I shall argue) that we can make sense of the expression of your resentment – its call – to be something like a report of the import both we and the circumstances have to us.⁹ Thus, in expressing your resentment, in issuing this report, you are in effect calling my attention both to the import to us of my norm violation and to your and my dignity as community members. Insofar as both we and these norms really do have this import to us and insofar as I am one of us, I ought therefore to respond to the import of my norm violation in the present circumstances with, for example, feelings and expressions of guilt. Consequently, on this account the call of the reactive attitudes operates only indirectly: to express a reactive attitude is to call another's attention to the relevant import, where it is this import itself that rationally motivates another's response.

This account of the call of the reactive attitudes cannot be right, for it ignores how that call depends on the authority of the one making it. As indicated earlier, outsiders, not being members of the relevant community of respect, are not in a position to feel such reactive attitudes and thereby to praise or blame members for notable upholdings or violations of the community's norms. Yet on the account just proposed, such authority is irrelevant, and *anyone* can have the same effect by making a

⁹ Something like this would seem to be the way content accounts of emotional expressions would have to understand the call of the reactive attitudes, at least if they took seriously the role of reactive attitudes in issuing that call.

report, since what rationally motivates the appropriate response is the import being reported rather than the report itself. Part of the point here can be put another way: this account ignores the status of the reactive attitudes as ways of *praising* or *blaming* the perpetrator and so of their expressions as ways of communicating that praise or blame. Something more seems to be involved in communicating praise or blame than just issuing a report concerning the upholding or violating of a norm. What more is this? As indicated earlier, Darwall's answer is to understand such blame in terms of the demand that the perpetrator respond by taking responsibility in some appropriate way, so that blaming presupposes the authority to make such a demand. Consequently, to express a reactive attitude, perhaps through such denunciation, reproach and so on, just is to make such a purportedly authoritative demand.¹⁰

Even so, we cannot understand the call of the reactive attitudes simply in terms of the authoritative demands issued in blame, for the call of the reactive attitudes can be directed not merely to the perpetrator but also potentially to the victim or to witnesses. For example, the victim's expression of resentment can call on witnesses to feel disapprobation or indignation and thereby to provide what we might call 'moral support'; a witness's expression of approbation can call on the victim to feel gratitude, and the perpetrator's expression of guilt can call on the victim to respond with at least an acknowledgement that the perpetrator has taken responsibility. Indeed, different ways of expressing a reactive attitude can target different parties with its call: these can be *targeted expressions*, and a particular expression can have multiple targets. Each of these cases involves an authoritative demand for recognition respect, and yet none of them involves blame. Hence it would be misleading to think that blame is at the centre of the call of reactive attitudes. Rather, what is essential to this call is the authority of the subject, the dignity she has as a member of the community of respect.

At this point we can see how the commitment account of emotional expression can explain this call of reactive attitudes. I said in [Section 4.2](#) that, from the objective perspective, emotions just are intentional feelings of import, a mode of attunement to import *impressing* itself on one

¹⁰ Indeed, this need for authority is a point that Tim Scanlon (2008) misses with his understanding of blame and blameworthiness in terms of the impairment of certain sorts of relationships between the blamer and the perpetrator: as I have argued elsewhere (Helm forthcoming), properly understanding which relationships are relevant here – something Scanlon leaves unspecified – requires bringing out the authority of the blamer in the context of a community of respect.

in feeling. In the case of the reactive attitudes, we can say something similar, albeit in the context of import we share jointly. Thus, when I interrupt your talk and so violate our norms as members of a philosophical community, the import to us of my doing so ought to impress itself on me by virtue of my attunement to such import as one of us, so that I ought to feel guilt, though of course I may not in particular cases or even in general: I may be inadequately attuned to the import this norm has to us. Yet, in addition, your expression of your resentment itself carries import to us, given your dignity as one of us, and I as a fellow member likewise ought to be attuned as one of us to this import so that in the circumstances this import also ought to impress itself on me in feeling; that this is so is a condition of the possibility of our being fellow members of a community of respect. This means that when you express your resentment, you are doing something to impress – you are *pressing* – that import on me, where the relevant import is the content of your expression, namely the import to us not only of you as a fellow member but also of my actions in the present circumstances. Hence the expression of a reactive attitude presses this import on fellow members by virtue of one's authority as a member. Indeed, to press import on a fellow member in this way provides a clearer understanding of what Darwall should mean in saying that (expressions of) reactive attitudes involve an authoritative 'demand', for the idea of a demand does not clearly fit many cases, especially those involving positive reactive attitudes such as approbation and gratitude. Consequently, by thus pressing import on fellow members, expressions of reactive attitudes are far from mere reports.

This same point emerges from the subjective perspective on caring. In general I have claimed that from the subjective perspective we understand an emotion to be a commitment both to the import of its focus and thereby to attend to its well-being and, when appropriate, act on its behalf; emotional expressions, then, are expressions of this commitment. In the case of the reactive attitudes, the relevant commitments are *our* commitments, and they are commitments to the members of the relevant community and its norms and thereby to attend and respond to the present circumstances when they affect the well-being of its members or norms. In particular, to be committed to the well-being of someone as a fellow member of a community of respect means being committed, other things being equal, to share in her commitments undertaken and expressed as a fellow member, for the failure in general to do this is to harm her status as a fellow

member. Indeed, each being normally thus committed to others as fellow members is a condition of the possibility of their forming a community of respect in the first place. Consequently, when you express a reactive attitude as a fellow member, I am thereby committed to attend and respond to this, so that it is my commitment as one of us that gives you a claim on my regard – that makes your expression of the reactive attitude be the expression of *our* commitment to the import of the relevant members and norms as this plays out in the circumstances.

This means that the *authority* of fellow members that is in view from the objective perspective is revealed from the subjective perspective to be the commitment each has as a fellow member both (1) to the import of the others as fellow members of a community of respect, and thereby (2) to share in their commitments undertaken as fellow members. Likewise, the *call* of another's reactive attitudes, which I understood from the objective perspective as the subject's authoritative pressing of the relevant import on a fellow member given the latter's corresponding attunement to that authority, is revealed from the subjective perspective to be the subject's expression of a commitment as *our* commitment, which gets a claim on fellow members insofar as they, each as one of us, are committed to her as a fellow member and so to being attentive and responsive to her expressions of our commitments.

In short, we can make the best sense of the call of the reactive attitudes and the role it plays in establishing and maintaining the jointness of values within a community of respect through the commitment account. Combined with the need, as I've argued, to distinguish between emotional expressions and reports and indirect corroboration from cases of actions directly expressive of import, this makes for a powerful case for the commitment account of emotional expression.

References

- Alston, W. P. 1965. 'Expressing', in Black (ed.) *Philosophy in America*. Ithaca, NY: Cornell University Press, pp. 15–34.
- Ayer, A. J. 1952. *Language, Truth and Logic*. New York: Dover.
- Benson, J. 1967. 'Emotion and Expression', *The Philosophical Review* 76: 335–357.
- Betzler, M. 2007. 'Making Sense of Actions Expressing Emotions', *Dialectica* 61: 447–466.
2009. 'Expressive Actions', *Inquiry* 52: 272–292.
- Brandom, R. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

- Darwall, S. L. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.
- Davis, W. A. 1988. 'Expression of Emotion', *American Philosophical Quarterly* 25: 279–291.
- Döring, S. A. 2003. 'Explaining Action by Emotion', *Philosophical Quarterly* 53: 214–230.
- Dretske, F. I. 1986. 'Misrepresentation', in Bogdan (ed.) *Belief: Form, Content, and Function*. Oxford: Oxford University Press, pp. 17–36.
- Ekman, P. 1993. 'Facial Expression and Emotion', *American Psychologist* 48: 384–392.
- Gibbard, A. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard University Press.
- Goldie, P. 2000. 'Explaining Expressions of Emotion', *Mind* 109: 25–38.
- Green, O. H. 1970. 'The Expression of Emotion', *Mind* 79: 551–568.
- Helm, B. W. 2001. *Emotional Reason: Deliberation, Motivation, and the Nature of Value*. Cambridge: Cambridge University Press.
2002. 'Felt Evaluations: A Theory of Pleasure and Pain', *American Philosophical Quarterly* 39: 13–30.
2008. 'Plural Agents', *Noûs* 42: 17–49.
2010. *Love, Friendship, and the Self: Intimacy, Identification, and the Social Nature of Persons*. Oxford: Oxford University Press.
2015. 'Rationality, Authority, and Bindingness: An Account of Communal Norms', in Shoemaker (ed.) *Oxford Studies in Agency and Responsibility*. Oxford: Oxford University Press, pp. 189–212.
- Forthcoming. 'Personal Relationships and Blame: Scanlon and the Reactive Attitudes', in Hutchison, Mackenzie, and Oshana (eds.) *Social Dimensions of Moral Responsibility*. Oxford University Press.
- Hursthouse, R. 1991. 'Arational Actions', *Journal of Philosophy* 88: 57–68.
- Koch, P. J. 1983. 'Expressing Emotion', *Pacific Philosophical Quarterly* 64: 176–189.
- Kovach, A. and De Lancey, C. 2005. 'On Emotions and the Explanation of Behavior', *Noûs* 39: 106–122.
- Scanlon, T. M. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, MA: Belknap Press of Harvard University Press.
- Strawson, P. F. 1962. 'Freedom and Resentment', *Proceedings of the British Academy* 48: 187–211.
- Teroni, F. 2007. 'Emotions and Formal Objects', *Dialectica* 61: 395–415.
- Watson, G. 2008. 'Responsibility and the Limits of Evil: Variations on a Strawsonian Theme', in McKenna and Russell (eds.) *Free Will and Reactive Attitudes: Perspectives on P. F. Strawson's 'Freedom and Resentment'*. Burlington, VT: Ashgate, pp. 115–141.

5. Collective Emotion and the Function of Expressive Behaviour

Carolyn Price

5.1 Introduction

It is uncontroversial that emotion has a strong social character. Our emotional responses are profoundly influenced by the behaviour and expectations of people around us. Moreover, many emotions have social functions: they enable us to sustain relationships with others and to cope with a range of social challenges (see Helm, this volume). More controversially, it is sometimes suggested that emotional responses can *themselves* be social phenomena: that emotions can sometimes be experienced *collectively* in some sense. Whether we can make interesting sense of this claim is a question that is attracting increasing attention from philosophers, reflecting both a flourishing literature in the philosophy of emotion and a vigorous debate about other collective phenomena – collective action and collective cognition, in particular.¹

In this chapter, I shall explore a new way of approaching this question. Rather than beginning by asking whether emotions themselves can be collective, I shall focus, first, on particular emotional *phenomena* – specifically, on emotional actions and on processes of reasoning and reflection that arise in the course of an emotional response. I want to consider whether these phenomena can sometimes occur collectively. As I shall explain, this is an interesting question in its own right; moreover, approaching the issue in this way suggests a new way of understanding the claim that emotions themselves can sometimes be collective. As we shall see, investigating these phenomena involves understanding

¹ For recent discussions of collective emotion, see Gilbert (2002, 2014); Schmid (2009); Huebner (2013); Salmela (2012); Slaby (2014).

the nature and function of expressive behaviour. Hence, a discussion of expressive behaviour will play a crucial role in what follows.

I begin, in [Section 5.2](#), by explaining why it might be thought that emotions *cannot* be collective in any interesting sense, and by briefly considering some possible responses to this concern. In [Section 5.3](#), I introduce the question that is my primary focus in this chapter: I explain why it is an interesting question, and how it relates to the broader debate about collective emotion. In [Section 5.4](#), I sketch an account of expressive behaviour; and in [Sections 5.5](#) and [5.6](#), I draw on this account in describing some potential cases of collective emotional action and reflection. I end, in [Section 5.7](#), by considering what conclusions might be drawn from these cases.

5.2 Approaches to Collective Emotion

People often talk as if there can be collective subjects of emotion: crowds become enraged; communities grieve; nations are proud of their history. Still, there are ways of understanding these claims without supposing that crowds or communities or nations are themselves subjects of emotion. The claim that a community is grieving, for example, can be understood merely as shorthand for the claim that most of its *members* are grieving. Indeed, the suggestion that there can be collective emotions in any more robust sense might well look doomed from the start. This is because (it is assumed) emotions are conscious mental states, with a particular phenomenal character; conscious states, it is widely supposed, can be attributed only to individuals.

There seems, then, to be a simple argument against the view that an emotion can be attributed to a group:

P1: An emotion is a conscious mental state, characterised by a certain kind of phenomenal character.

P2: Only individuals can have conscious mental states.

Conclusion: So, only individuals can have emotions (emotional individualism).

In this section, I shall review, very briefly, three possible responses to this argument. Because I need to be brief, I cannot do justice to the details of the views mentioned here; hence, there remains much more to be said about the issues raised in this section. My aim is just to provide some background to the discussion that follows and to motivate a search for an alternative approach.

The simple argument is not decisive: both premises are open to question. Consider, for example, Margaret Gilbert's well-known account of collective guilt. According to Gilbert (2002, 2014), collective guilt arises when a group of people (a board of directors, say) are jointly committed to behaving in certain ways – apologising, making reparations – when they agree that they have made some poor collective decision. As she acknowledges, they might do this without experiencing any feelings ('pangs or twinges') of guilt. Still, she argues, this is not an objection to her account. Emotions, she says, do not always involve feelings; in other words, they do not always have a particular phenomenal character (Gilbert 2002: 119–120). Hence, Gilbert seems to reject P1.

However, it is not clear that P1 can plausibly be defeated in this way. As many theorists of emotion have pointed out, the phenomenal character of an emotional response is not easily viewed as a mere accompaniment of emotion.² Rather it reflects key features of the response as a whole. These include not only changes taking place in the subject's body but also a range of psychological changes: the subject's intense attentional focus on the situation; how the scene looks and sounds; the thoughts racing through their head; the urgent desire to act. Were many of these features missing, it is hard to say why the response should be viewed as an *emotional* response. Plausibly, then, the absence of phenomenology is not merely the absence of an optional extra: it puts serious pressure on the claim that collective guilt, as Gilbert describes it, is a type of emotion.³

Another option might be to reject P2. This option has recently been investigated by Bryce Huebner (2013). It is at least conceivable, Huebner (2013: 112) suggests, that we might one day encounter a collective so vast and tightly organised that it is able to match the computational complexity of the human brain; if so, we could have reason to believe that it was capable of subjective experience. Arguably, then, P2 should be more cautiously phrased. Still, as Huebner also points out, there is no reason to think that groups of the kind we ordinarily encounter (boards of directors, teams of scientists) are sufficiently large or complex to generate conscious experience. Hence, assuming

² See, for example, Goldie (2000: 50–83); Montague (2006); Helm (2009). The worry that Gilbert underestimates the importance of emotional phenomenology has been voiced by a number of commentators: see particularly Wilkins (2002); Konzelmann Ziv (2009).

³ Some theorists might reject P1 on the grounds that bodily changes, rather than phenomenology, are central to emotion (see Sias and Bar-On, this volume). This move, though, would not help Gilbert.

P1 is correct, it looks as if collective emotion remains at best a theoretical possibility.⁴

In contrast, Mikko Salmela (2012) does not try to challenge the simple argument. Instead, he sets out to develop a robust notion of *shared* emotion, entirely compatible with emotional individualism (Salmela 2012: 44). Cases of shared emotion, he suggests, involve four elements:

1. The emotional responses of those involved are grounded in the same concern (e.g., the success of a football club).
2. The individuals involved have the same type of emotional response and undergo similar physiological, psychological and behavioural changes.
3. Mechanisms of attentional deployment, emotional contagion, facial mimicry and behavioural entrainment help to regulate and intensify the emotional responses of the individuals involved.
4. The individuals involved are aware that they share the same emotion. (Salmela 2012: 39–42)

This concept of shared emotion, Salmela points out, is a matter of degree. In the strongest cases, he suggests, those involved are collectively *committed* to the concern that grounds the emotion; the concern is itself collective ('our jointly achieved success'); and the emotional response is relatively intense. An example might be the mutually contagious joy of the members of a football team as they stand together on the pitch having just won an important trophy for the club they all love (Salmela 2012: 42–43).

Salmela, then, offers a middle path: his concept of shared emotion goes well beyond the boring thought that two people can experience the same *type* of emotion; nevertheless, his account is perfectly compatible with emotional individualism. The idea that there is room for a middle path seems right; moreover, Salmela's account looks promising. Still, there remains scope to question whether his account of shared emotion could be extended or improved: in particular, we might ask whether the four elements he identifies are the right ones. I shall return to this question at the end.

There remains, too, a more fundamental question. Should we now rest content with (something like) Salmela's notion? Or are there

⁴ However, Huebner (2013: 244) also seems to suggest that a group such as an election campaign team can, in certain circumstances, be anxious or afraid. Presumably, then, he also rejects P1.

possibilities yet to explore – possibilities that might loosen the grip of emotional individualism? In what follows, I shall pursue this further question: in particular, I shall focus on the possibility that there are collective emotional *phenomena*.⁵ In the next section, I shall explain what I mean by this, and why I think that this possibility is worth investigating.

5.3 Collective Emotional Phenomena?

Getting emotional is complicated: it might involve changes to heart rate and skin conductance; changes to attention and motivation; memories, imaginings and ruminations; blushing, frowning, sighing, protesting, celebrating, or fleeing a bear. Characteristically, these changes are generated by an evaluation of the situation; and, together, they generate a conscious emotional experience, with a specific phenomenal character. When I use the term ‘emotion phenomena’ in what follows, it is these changes – physiological, psychological or behavioural – that I have in mind. Emotional phenomena, then, include emotional evaluations, emotional feelings, emotional desires, emotional expressions, emotional actions and so on. Together, these changes form a complex emotional response – for example, an episode of fear (Price 2012, 2015).⁶

To describe something as an emotional phenomenon is not to identify it as an emotion or even as a component of an emotion. Some emotional phenomena might be classed as causes or effects of the emotion. Hence, the claim that there can be collective emotional phenomena is not equivalent to the claim that there can be collective emotions, as such. Nevertheless, it is an interesting claim in its own right. Suppose, for example, that it could be shown that some emotional actions are collective actions: such a finding would almost certainly have significant implications for the kinds of intentional content that emotional evaluations can have, for the kinds of emotional motivations there can be, and for the functions of other emotional phenomena, such as expressive behaviour. Moreover, we might well want to take account

⁵ In what follows, I use the term ‘shared’ when I am referring to a notion that, like Salmela’s, is compatible with emotional individualism; I use the term ‘collective’ when I am referring to the idea that emotions or emotional phenomena can sometimes be attributed to groups.

⁶ Elsewhere (Price 2012, 2015), I use the term ‘emotional phenomenon’ to refer to complex emotional responses too. I am using the term a little more narrowly here.

of this finding in explaining what it is for an emotion to be *shared*, in Salmela's sense.

Moreover, such a finding may yet have a bearing on whether emotions themselves can be collective. To explain why, I need to distinguish some different ways of answering the question 'What is an emotion?'

1. *Single component theories*. Some theorists identify emotions with a *single* emotional phenomenon. William James (1890) famously held that emotions are bodily feelings. More recently, many theorists have taken emotions to be emotional evaluations (though they disagree about what these evaluations are).⁷
2. *Complex process theories*. Peter Goldie (2000: 12–14) argues that an emotion is a complex *process*, which has many components, including perceptions, thoughts, feelings, physiological changes and a variety of emotional and behavioural dispositions. Jenefer Robinson (2005: 57–61) also holds that an emotion is a process in which an affective appraisal triggers physiological and motor changes, action tendencies, changes in vocal and facial expression and subsequent cognitive monitoring of the situation.⁸ Among psychologists, Paul Ekman (1982, 1992) has long argued that emotions are complex, organised responses involving a range of physiological, behavioural and psychological changes.
3. *An adjectivalist account*. My own view is different again. As I have explained elsewhere (Price 2012, 2015), I am not convinced that the question 'What is an emotion?' *can* be settled in any satisfying way. Hence, I prefer to do my theorising about emotional *phenomena* – emotional evaluations, feelings, actions and so on – and about the complex emotional responses that they compose. We can study these things, I want to say, without needing to decide which of them constitutes *the* emotion.

Focus, first, on complex process theories. These theories tend to differ from single component theories in one crucial respect: while single component theorists tend to identify emotions with some kind of mental state (feelings or evaluations), complex process theorists generally allow that emotions include at least some non-psychological components. Admittedly, there is some disagreement about what the

⁷ This view is endorsed by Solomon (1993); Helm (2009); Roberts (2003); Prinz (2004), among many others.

⁸ Compare Salmela (2012: 42).

components of an emotion are. As I mentioned, Robinson, Salmela and Ekman include certain kinds of expressive behaviour, while expressive behaviour is explicitly excluded by Goldie (2000: 13).⁹ Still, they all agree that emotions include at least *some* non-psychological phenomena. Hence, none of them is committed to P1: they hold that emotions include a *mix* of psychological and non-psychological components. As a result, none of them should endorse the simple argument for emotional individualism.

Still, this does not entail that they should reject emotional individualism as such: it may yet turn out that emotions are necessarily confined to individual bodies, if not to individual minds. Moreover, they all agree that emotions have some psychological components: hence, they cannot suppose that an emotion could be *wholly* collective. Nevertheless, there remains an interesting possibility – that emotions can be *partly* collective. Suppose, for example, there is some phenomenon – emotional reflection, say – that a complex process theorist normally regards as a component of emotion. Suppose, too, that it is established that emotional reflection sometimes occurs within groups, rather than individuals. This raises the possibility that such a process of collective reflection might itself be a component of an emotion.

At first blush, this suggestion certainly sounds odd. However, it sounds less odd if we take seriously the idea that emotions are processes. Processes can evolve in all kinds of ways: they can start in one place and end in another; they can involve different objects or people at different times; they can overlap, in part, with other processes going on at the same time. Suppose, for example, that two neighbours, Babs and Bob, are both busy tidying their gardens: most of the time they work separately, but they work together to trim the dividing hedge. Arguably, there are two (partly overlapping) processes here – Bob tidying his garden and Babs tidying hers – each of which includes their trimming the hedge together. The suggestion is that emotions, too, might sometimes include a collective stage.

For me, as an adjectivalist, the issue is slightly different. Because I take no particular view about what an emotion is, the question

⁹ For Ekman (1992) and Robinson (2005), the emotion consists of all the changes that are automatically triggered by the emotional appraisal: this, they think, includes non-voluntary expressive behaviour. I am not altogether clear why Goldie draws the line exactly where he does. The thought may simply be that we intuitively regard any kind of emotional behaviour as caused by, rather than part of, the emotion. For a concern about this approach, see Price (2012: 326–328).

whether there can be collective emotions, as such, does not arise. Nevertheless, as I mentioned earlier, I am certainly interested in understanding emotional *responses*. An emotional response, I take it, is a complex process: it has a particular function and structure; and it includes both psychological and non-psychological components (Price 2015: 2–8). Moreover, as I am using the term, an emotional response includes *all* the emotional phenomena that the subject produces or undergoes. These will include emotional phenomena that complex process theorists do not tend to regard as components of emotions. Consider, for example, non-expressive emotional actions (fleeing a bear or throwing a punch, say): none of the complex process theorists mentioned lists these emotional actions as components of emotions. Nevertheless, such actions will certainly be part of the broader emotional response. Hence, the discovery that there can be collective emotional phenomena would certainly raise the possibility that emotional *responses* are sometimes partly collective – whatever those phenomena turn out to be.

Can there be collective emotional phenomena? As will become clear in Sections 5.5 and 5.6, my answer to this question depends on some particular claims about the function of expressive behaviour. My next task, then, is to make those claims.

5.4 Expressive Behaviour

The term ‘expressive behaviour’ has been understood in many ways in the philosophical literature.¹⁰ Here, I shall explain how I understand the term.

We might distinguish between three kinds of emotional behaviour:

1. Behaviour aimed at dealing with the particular challenge or opportunity that has elicited the emotion. This might include angrily thumping an adversary, fearfully fleeing a bear, guiltily buying flowers. It also includes certain kinds of cognitive behaviour: anxiously fretting about a problem, triumphantly gloating over a victory.
2. Behaviour aimed at managing one’s emotional response. This might include venting your frustration by swearing; blocking your ears to shut out the awful sound; working yourself up into a rage.
3. Expressive behaviour. This might include blushing; frowning; whimpering; turning a cartwheel; saying ‘Oh God, it’s coming this way!’; stroking someone’s face; wearing a black armband; sending flowers.

¹⁰ For some recent discussions, see Davies (1988); Goldie (2000); Green (2007).

Emotional behaviour often belongs to more than one of these categories. As I list the crimes of an adversary to a mutual acquaintance, I may be expressing my resentment, amplifying it and getting my revenge all at the same time. In particular, expressing an emotion is often a way of resolving the situation: expressing anger, for example, can be a way of deterring further offence; expressing love can be a way of cementing a bond. Nevertheless, not *all* emotional behaviour is expressive: fleeing a bear, for example, manifests fear, but it would seem odd to say that it *expresses* it.

Like other kinds of emotional behaviour, expressive behaviour is not all of a kind. Some of the examples listed are automatic, involuntary behaviours, while others are intentional actions. Some seem to be hard-wired, while others presuppose a sophisticated background of cultural conventions, including linguistic or symbolic conventions. In other cases, expressive behaviour can be quite personal, even idiosyncratic: expressing one's rage by planting tulips, say.

Given the diversity of expressive behaviours, we might wonder what they have in common. First, expressive behaviour is itself *emotional* behaviour: it is either an involuntary response triggered by an emotional evaluation or a voluntary action motivated by an emotional desire. Hence, a 'thank you' present expresses my gratitude only if I give it out of gratitude – and not, say, out of duty. For the same reason, there is a difference between expressing emotion and merely reporting it. As we have seen, though, not all kinds of emotional behaviour count as expressive behaviour: fleeing a bear, for example, is not expressive behaviour. A natural explanation for this is that describing behaviour as expressive implies that it has a *meaning*: expressive behaviour is supposed to signal or convey how the subject feels. This might be taken to imply that expressive behaviour always has a *communicative* function: it functions to communicate one's emotional response to others.¹¹ This certainly seems right in some cases. However, there are two qualifications to make.

First, expressive behaviour does not always express the subject's emotional response *overall*: some expressive behaviour is concerned only with a specific component of the response. Crying out 'Oh God, it's coming this way!' might be described as expressing fear; more

¹¹ Here, I am using the term 'communication' in a broad sense to include both intentional acts of communication (e.g., telling someone the time) and non-intentional behaviour (e.g., bees' waggle dances). For a detailed account of expressive behaviour as communicative in character, see Green (2007).

precisely, though, it expresses a particular fearful *thought*. Similarly, nervously tapping one's watch might express an anxious desire to leave. This point is particularly important in accounting for expressive behaviour that is symbolic in character. Goldie (2000: 134–135) describes a husband angrily smashing his wife's favourite vase in front of her. One possibility, he suggests, is that his smashing the vase expresses not simply his anger, but his angry desire to harm her: it does this by symbolising the harm he wants to do. On another scenario, smashing the vase might express an angry thought – for example, that the marriage is broken beyond repair.

Second, it can be questioned whether all expressive behaviour is communicative behaviour. Consider Hursthouse's (1991) example of angry Jane who grabs a photograph of her rival, Joan, and scratches out the eyes. Jane's behaviour does not seem to have a communicative character: we are not supposed to imagine that Jane damages the photograph in front of Joan – or anyone else. Goldie (2000: 130) suggests that Jane's behaviour can be understood as a case of wish fulfilment. Understood in this way, her behaviour has no particular function: it simply reflects the brute psychological fact that when people cannot fulfil some strong desire, they sometimes take satisfaction in symbolically acting it out. Nevertheless, Goldie classes this as a case of expressive behaviour: in symbolically acting out her wish, he suggests, Jane thereby expresses it (Goldie 2000: 129–134).¹²

Is this right? I agree that Jane's behaviour might well be a case of wish fulfilment; but, if so, I am not sure that her behaviour should be said to *express* her wish. Earlier, I suggested that expressive behaviour is naturally viewed as behaviour that is supposed to signal or convey the subject's emotional response. If Jane is simply acting out her wish for her own satisfaction, her behaviour certainly manifests her wish; but it does not seem to express it, in this sense.

However, there is another possible explanation for Jane's behaviour – one that makes better sense of the idea that this is expressive behaviour. Sometimes, expressing a wish is not a way of communicating it to others, but rather a way of articulating it to oneself. Craving for chocolate, for example, I mutter: 'I wish I had some chocolate right now!' My utterance need not be aimed at anyone else: it might just be a way of *articulating* my desire – of holding it in consciousness. Similarly, I might close my eyes and imagine the taste, not because this satisfies

¹² Compare Green (2007: 36–37).

my desire (far from it), but just as a way of attending to it: 'It is *this!*', I say to myself, 'It is *this* that I crave!' Conceivably, this might be what Jane is doing when she attacks the photograph: '*This!*' she says, '*This* is what I want to do to *you!*' On this scenario, Jane's behaviour can be regarded as expressing her desire – not because her behaviour has a communicative function, but because it functions to *articulate* her emotional desire.¹³ In other cases, expressive behaviour might articulate an emotional *thought*: by casting the photograph aside, for example, Jane might articulate the thought that Joan is someone of no consequence. (Compare Bennett's suggestion, this volume, that expressive behaviour may be aimed at acknowledging or reflecting what is significant about the situation.)

I have suggested that expressive behaviour is highly diverse. It includes both automatic responses and intentional actions; it embraces hard-wired behaviours, conventional gestures and idiosyncratic performances, interpretable only by one's nearest and dearest. What these cases have in common is (1) that they are cases of emotional behaviour; and (2) that they function to communicate or articulate something about the subject's emotional response – either the subject's response overall, or a specific emotional thought or desire. In the next two sections, I shall describe two possible kinds of collective emotional phenomenon: collective emotional action and collective emotional cognition, drawing on the account of expressive behaviour sketched here.

5.5 Collective Emotional Action

Henry is a technician in a medium-sized company. Year on year, the managers have awarded themselves large bonuses, citing rising profits. This morning, Henry's manager, Sandra, has announced that, for the third year running, none of the technical staff will receive a rise. Having read Sandra's memo, Henry looks round at his colleagues and sees expressions of disbelief and anger on their faces. One flings herself angrily back in her seat, while another slams his hand on his desk. Angry mutterings develop into animated discussion. Eventually, action emerges: the technicians march together to Sandra's office to demand that the decision is reversed.

On the face of it, this looks very much like a case of collective action. Admittedly, there is some dispute about what exactly is required for

¹³ Compare Slaby (2014: 36).

genuine collective action. Still, the current consensus seems to be that we can answer this question without attributing beliefs or intentions to the group itself, but rather by appealing to the goals, perceptions and dispositions of its individual members.¹⁴ In this case, the following points all seem significant:

1. Henry and his colleagues each act with roughly the same goal (getting the decision reversed).
2. Each is disposed to accept the others' help, and to coordinate their behaviour with the others, in order to achieve this goal.
3. Each is aware that the members of the group share this goal and these dispositions.

Given all this, I would suggest, it makes sense to say not just that each individual protested, but also that the group itself protested *as a body*.

However, to establish that this is a case of collective action is not to establish that this is a case of collective *emotional* action. Before we can settle this there are two further hurdles to overcome.

First, for an action to count as an emotional action, collective or otherwise, it must be done out of emotion: it must be motivated by emotional desires or urges. The technicians' indignant response certainly involves an urge to protest; but it is not clear that we can appeal to this urge to explain why they protested *together*. It might be thought, rather, that the situation must have developed in something like the following way: having read Sandra's memo, each technician was seized by an indignant urge to go and personally remonstrate with Sandra; on reflection, however, each realised that they could achieve nothing by acting alone; as a result, they each *restrained* their immediate impulse to confront Sandra, and instead consulted each other – eventually agreeing to mount their protest together. Hence (the worry goes) when the technicians decide to protest *as a body*, they are not acting on their initial, indignant urge, but rather on a more considered desire to ensure that their voices are heard.

However, there is another possibility. Suppose that there is a particular *type* of righteous anger – one that motivates the subject, not merely to resist the wrong, but to resist it *in concert with others*. (In what follows, I shall commandeer the term 'outrage' as a label for this form of anger.) The suggestion that there is such a type of anger strikes me as quite

¹⁴ For some recent accounts of this kind, see Bratman (1993); Tuomela (2005); Tollefsen (2005); Isaacs (2011).

plausible. Certainly, it is easy to see how people would *benefit* by having an emotional propensity of this kind: when a perpetrator is too powerful, or insouciant, to be tackled by one person acting alone, people would do well to be motivated to act together with others. Moreover, angry people do often seem to have an urge to broadcast their wrongs – heatedly drawing others’ attention to the situation and trying to gather others to their cause. Indeed, it seems plausible that there are other emotional responses of this kind: triumphant joy, for example, often seems to move people to celebrate *with* others; and it is not hard to imagine how people might benefit from this impulse. Suppose, then, that the technicians respond to Sandra’s memo with outrage, in this sense. If so, it is possible that what motivated their joint protest is not a considered desire to be heard, but an outraged urge to act together. Hence, assuming that this is a real possibility, we can make room for the idea that the technicians protest together out of emotion.

There is a second hurdle to jump. The technicians, we suppose, share an outraged urge, not merely to protest, but to protest together. Still, they might not yet be in a position to act *effectively* on this motivation: they may first need to decide exactly what to do. This might well require an explicit process of planning and coordination. Once again, the worry will be that this will require them to *restrain* their outrage, so that they can discuss calmly how to organise themselves. But if so, this looks problematic: for on this scenario, it is not altogether clear that their protest, when they mount it, is an *emotional* action, or at least not a purely emotional action. Even if their protest was initially motivated by an outraged urge to protest – and even if this urge continues to strengthen their resolve – nevertheless, by the time they actually mount their protest, they are not acting directly on this urge, but rather on a plan that has been formulated, and endorsed, in cold blood. Matters would be different, it seems, if we could suppose that their planning is not calm at all, but emotional – directed by outraged thoughts, imaginings and desires – just as someone might angrily plan an act of personal revenge. If so, it would be much clearer that their protest is an *outraged* protest – motivated, planned and performed wholly in the grip of emotion.

How, though, could this come about? The answer, I take it, lies with expressive behaviour: for it is through expressive behaviour that the technicians are able to communicate their outraged thoughts and desires – and to do this heatedly, rather than calming reporting their emotional state. John Michael (2011) has already offered a helpful

analysis of the ways in which expressive behaviour might facilitate collective action. By allowing us to recognize others' emotions, Michael suggests, expressive behaviour helps each individual to predict what the others are likely to do; and the emotional rapport shared by the members of a group may lead them, quite unconsciously, to move in alignment. Michael's concerns, however, are not quite the same as mine: he is concerned with collective action in general, while I am concerned only with collective *emotional* action. As a result, my account departs (at least in emphasis) from his in two specific ways.

First, I want to stress the importance of verbal or symbolic expressions of specific emotional thoughts and desires. This is because I want to allow that there can be collective emotional actions of a relatively sophisticated kind. Almost certainly, planning and coordinating these actions will require the participants to communicate, not just the broad tenor of their emotional response, but specific emotional thoughts, imaginings and desires. (These might include, e.g., thoughts about what might make for a particularly telling protest, or the desire to get others involved.) Second, Michael is chiefly concerned with the way in which expressive behaviour facilitates collective action. In contrast, I want to stress that, when it comes to collective emotional action, the role of expressive behaviour is not only causal, but also constitutive. It is partly *because* the technicians' protest was planned in the grip of outrage – planned, that is, in response to behaviour that expressed outraged thoughts and desires – that it counts as a clear case of collective *emotional* action.

I have argued that some cases of collective action might turn out to be cases of collective emotional action. This will be possible provided that there are some types of emotional response that motivate people to act together. Moreover, if we are to allow for collective emotional actions of any sophistication, we must also allow that collective actions can be planned and coordinated in an emotional way. This is possible, I have suggested, because people can employ expressive behaviour in planning actions of these kinds.

5.6 Collective Emotional Cognition

Emotional responses frequently involve processes of thought or reasoning. Most obviously, they involve processes of practical reasoning: in anxiety or anger, people cast around for a way to avert the catastrophe or redress the wrong. Other types of emotional response characteristically

involve an element of reflection or rumination: in grief, the subject is often moved to reflect on their loss – to try to understand what they have lost and why (Price 2010). Anxious fretting and sorrowful reflection are often solitary activities. We might wonder, though, whether emotional cognition can sometimes be collective.

Questions about collective cognition have often been pursued within the context of a wider debate about collective intentionality: it is often thought that to describe a cognitive process as collective implies ascribing beliefs or intentions to groups.¹⁵ However, this is not the only way in which collective cognition can be understood. Ronald Giere (2002) has argued that we can understand collective cognition without positing collective psychological states: the only psychological states required are those of the individual members of the group. Indeed, we might try to model collective cognition much as we modelled collective action in the last section: the idea will be that people can work together to solve a problem or to form a plan, just as they can work together to mount a protest.

With that in mind, it is not hard to find possible examples of collective emotional cognition. Consider, first, the case of Henry and his colleagues: to act together, they must coordinate their behaviour, and – as we have seen – this may involve a degree of explicit planning. As I suggested in the previous section, to the extent that this involves the expression of emotional desires and thoughts, this will be a case of *emotional planning*. Arguably, there can be cases of collective reflection, too. Consider the case of grief: in many cultures, rituals of mourning encourage the bereaved to reflect together – for example, by exchanging memories of the deceased. Exchanging memories can be comforting, but arguably, it also has an epistemic function: by reflecting together, the mourners are able to reach a shared understanding of their loss.

In the case just described, the mourners' expressive behaviour plays a communicative role: it enables them to communicate their sad thoughts. In other cases, it seems more natural to think of the process as functioning to *articulate* a certain way of understanding what has occurred. This makes better sense of cases in which mourners come together to perform some ritual action. Suppose, for example, that a community has come together to build a cairn for the deceased – expressing, perhaps, some sad but consoling thought. As each individual solemnly adds a stone to the cairn, it would seem strange to think of their behaviour

¹⁵ For some discussions, see Hutchins (1995); List (2008); Huebner (2013).

as communicating this thought to the others – each contributing some piece of information that the others lack. Rather it seems more natural to think of their behaviour as collectively highlighting, or even endorsing, that thought: together, they are *framing* the situation in a certain way. (Again, compare Bennett, this volume.¹⁶)

I have suggested, then, that we can make room for the idea of collective emotional cognition without positing collective psychological states. All these cases essentially involve expressive behaviour. Indeed, understanding the role of expressive behaviour in both communicating and articulating emotional thoughts and desires can help us to recognize the different forms that collective emotional cognition might take.

5.7 Shared and Collective Emotion

In the last two sections, I have identified two possible kinds of collective emotional phenomena: collective emotional action and collective emotional cognition. To do this, I have stressed, involves taking account of our capacity to express emotion in diverse and sometimes nuanced ways. I shall end by considering what further conclusions we might draw.

I shall start with Salmela's notion of shared emotion. As we saw, Salmela takes this to involve four elements: sameness of concern, sameness of response, mutual regulation and mutual awareness. He does not, though, accord a role to emotional action or cognition. I would like to suggest that these phenomena also have a role to play in cases of shared emotion. Consider Salmela's example of a team's joy at winning a trophy together. Imagine, first, that the players are standing together, aware of each other's feelings, and physically attuned to each other, just as he describes; moreover, each player is excitedly thinking about the team's victory and engaging in individual acts of celebration – dancing about and punching the air. Contrast this with a case in which the players are doing all these things, but also celebrating *together* – hugging each other and dancing around as a group; imagine, too, that they are excitedly *discussing* their victory, creating a shared understanding of their achievement. Plausibly, the second case represents a significantly more robust example of shared joy than the first. There is scope, then,

¹⁶ Indeed, it was hearing an earlier version of Bennett's paper at a workshop on 'Emotion and Expression' at the University of Manchester in 2012 that drew my attention to this kind of case.

to extend Salmela's analysis to take account of these further aspects of an emotional response. Moreover, this might well involve reconsidering the role of expressive behaviour in cases of shared emotion. For, as we have seen, collective emotional action and cognition characteristically depend not only on the sub-personal processes emphasised by Salmela, but also on the verbal or symbolic expression of specific emotional desires and thoughts.

In Section 5.3, I raised a further possibility: that complex emotional responses, perhaps even emotions themselves, might sometimes include a collective stage. Consider, first, an emotional *response* – an episode of outrage, say. The thought will be that, just as an episode of indignation might include some individual act of protest, so Henry's outraged response includes the collective protest mounted by Henry and his colleagues.

It might be objected that this is not the right thing to say. This is because there is a more modest alternative available: Henry's outraged response, we should say, includes, not the collective protest, but only Henry's individual *participation* in the protest. However, there is scope to resist this suggestion. To do this, we need to take seriously the idea that Henry's outraged response is a process, which has a particular function, and which is supposed to unfold in a particular way. The function of Henry's outrage is not, presumably, to ensure that he *participates* in redressing the wrong, but to ensure that the wrong is redressed – something Henry cannot achieve alone. Thus, what would *satisfy* Henry's outrage – what would bring it to its proper conclusion – is not Henry's *participation* in redressing the wrong,¹⁷ but – more simply – the fact that he and his colleagues got the decision reversed. Similarly, in a case of collective reflection, it is not my *participating* in some collective act of reflection that helps me to move on in my grief, but – more simply – the fact that we have found, together, a way to make sense of our loss. The thought is, then, that if we want to understand how these emotional responses unfold and how they are supposed to be resolved, ultimately, we need to view these collective phenomena themselves as stages in the process.

This is as far as an adjectivalist can go. Might a complex process theorist say more? As I explained earlier, it is open to a complex process theorist to claim that emotions themselves might sometimes be partly collective. They might have reason to do this, I suggested, if it

¹⁷ He might, of course, feel proud of his personal contribution.

turned out that some phenomenon that they would normally regard as a component of emotion can sometimes be attributed to a group. On this criterion, the possibility of collective emotional actions looks like a red herring; for, as I explained earlier, complex process theorists do not generally view non-expressive emotional actions as components of emotions.¹⁸ Cases of collective emotional cognition, however, hold more promise. Goldie, at least, suggests that thoughts are components of emotions (Goldie 2000: 12); and it seems a small step from there to suppose that emotions include *processes* of thought. Indeed, if an emotion is itself a process, it might make more sense to view its cognitive components as processes, rather than discrete mental acts. But if a complex process theorist regards processes of reasoning or reflection as components of the emotion, then they have some reason to allow that emotions might include collective reasoning and reflection too.

It might be objected, perhaps, that the two cases are not as closely analogous as I am implying. In particular, individual processes of reasoning and reflection are internal psychological processes, involving only the thoughts of the individual subject. Collective reasoning, in contrast, is a public, social phenomenon: it involves people saying or doing certain things, and perceiving and understanding each other's behaviour. Still, it is not clear that the contrast is as stark as this implies. After all, individual cognition is not always confined to the head: people sometimes think out loud; they solve problems by manipulating objects and symbols. Moreover, as I have stressed, complex process theorists *already* take emotions to include non-psychological components. Hence, a complex process theorist need not be committed to the view that the cognitive components of emotions are always purely internal psychological processes, even in cases involving a single subject. If so, this suggested disanalogy between individual and collective emotional cognition is not as significant as it might seem.

In this chapter, I have identified two kinds of collective emotional phenomena – collective emotional actions and collective emotional cognition; I have argued that these phenomena depend, in part, on our capacity for expressive behaviour. These collective emotional phenomena, I take it, are significant in their own right; but they also point to some further conclusions. In particular, I have suggested that Salmela's account of shared emotion should be extended to take more account

¹⁸ See note 9. As I indicated there, there is room for debate about this; but this is not an issue I can pursue here.

of emotional action and cognition. I have also suggested that a good case can be made for the claim that emotional responses can sometimes have a collective component or stage; moreover, on a complex process theory, a case could perhaps be made for emotions themselves. If so, expressive behaviour has the power to extend, not only the influence of emotion, but its boundaries too.

References

- Bratman, M. 1993. 'Shared Intention', *Ethics* 104: 97–113.
- Davies, W. 1988. 'Expression of Emotion', *American Philosophical Quarterly* 25: 279–291.
- Ekman P. 1982. 'Biological and Cultural Contributions to Body and Facial Movement in the Expression of Emotions', in Rorty (ed.) *Explaining Emotions*. Berkeley: University of California Press, pp. 73–102.
1992. 'An Argument for Basic Emotions', *Cognition and Emotion* 6: 169–200.
- Giere, R. 2002. 'Distributed Cognition in Epistemic Cultures', *Philosophy of Science* 69: 637–644.
- Gilbert, M. 2002. 'Collective Guilt and Collective Guilt Feelings', *The Journal of Ethics* 6: 115–143.
2014. 'How We Feel: Understanding Everyday Collective Emotion Ascription', in von Scheve and Salmela (eds.) *Collective Emotions*. Oxford: Oxford University Press, pp. 17–31.
- Goldie, P. 2000. *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press.
- Green, M. 2007. *Self-Expression*. Oxford: Oxford University Press.
- Helm, B. 2009. 'Emotions as Evaluative Feelings', *Emotion Review* 1: 248–255.
- Huebner, B. 2013. *Macro-cognition: Distributed Minds and Collective Intentionality*. New York: Oxford University Press.
- Hursthouse, R. 1991. 'Arational Actions', *Journal of Philosophy* 88: 57–68.
- Hutchins, E. 1995. *Cognition in the Wild*. Cambridge MA: MIT Press.
- Isaacs, T. 2011. *Moral Responsibility in Collective Contexts*. New York: Oxford University Press.
- James, W. 1890. *The Principles of Psychology*. New York: Dover, 1950.
- Konzelmann Ziv, A. 2009. 'The Semantics of Shared Emotion', *Universitas Philosophica* 52: 81–106.
- List, C. 2008. 'Distributed Cognition: A Perspective from Social Choice Theory', in Albert, Schmidtchen and Voigt (eds.) *Scientific Competition: Theory and Policy*. Tübingen: Mohr Siebeck, pp. 285–308.
- Michael, J. 2011. 'Shared Emotions and Joint Action', *Review of Philosophical Psychology* 2: 355–373.
- Montague, M. 2009. 'The Logic, Intentionality and Phenomenology of Emotion', *Philosophical Studies* 145: 171–192.
- Price, C. 2010. 'The Rationality of Grief', *Inquiry* 53(1): 20–40.
2012. 'Doing without Emotions', *Pacific Philosophical Quarterly* 93: 317–337.

2015. *Emotion*. Cambridge: Polity Press.
- Prinz, J. 2004. *Gut Reactions: A Perceptual Theory of Emotion*. Oxford: Oxford University Press.
- Roberts, R. 2003. *Emotions: An Essay in Aid of Moral Psychology*. Cambridge: Cambridge University Press.
- Robinson, J. 2005. *Deeper than Reason: Emotion and Its Role in Literature, Music and Art*. Oxford: Oxford University Press.
- Salmela, M. 2012. 'Shared Emotions', *Philosophical Explorations* 15: 33–46.
- Schmid, H. 2009. *Plural Action: Essays in Philosophy and Social Science*. Hamburg: Springer.
- Slaby, J. 2014. 'Emotions and the Extended Mind', in von Scheve and Salmela (eds.) *Collective Emotions*. Oxford: Oxford University Press, pp. 32–46.
- Solomon, R. 1993. *The Passions: Emotions and the Meaning of Life*. Indianapolis: Hackett.
- Tollefsen, D. 2005 'Let's Pretend! Children and Joint Action', *Philosophy of the Social Sciences* 35: 75–97.
- Tuomela, R. 2005. 'We-Intentions Revisited', *Philosophical Studies* 125: 327–369.
- Wilkins, B. 2002. 'Joint Commitments', *Journal of Ethics* 6: 145–155.

PART II

Psychological Perspectives

6. Emotional Expression

Hyisung C. Hwang and David Matsumoto

6.1 Emotions

6.1.1 Understanding Emotions

According to the *Oxford English Dictionary*, emotion is defined as ‘a natural instinctive state of mind deriving from one’s circumstances, mood, or relationships with others’. This standard definition implies interesting characteristics about the nature of emotions because emotions involve something innate, something that we are born with. Emotions are often invoked during social interactions and aid in communicating with others by sending and receiving messages through their expressions, which can vary depending on context. Of course, many scholars have their own definition of emotions and there are many differences among those definitions.

For this chapter we define one type of emotion as transient, biopsych-social reactions to events that have consequences for our welfare and potentially require immediate actions (Matsumoto and Hwang 2012). These emotions are biologically resident and are products of our evolutionary history, providing the platform for universality in the domain of immediate reactions (Matsumoto and Juang 2013). Our definition of emotion is based on the study of priming reactions that precede potential immediate action and that are tied to physiological responses, expressive behaviours and cognitive gating; this type of emotion is called ‘basic emotions’ (Ekman 1972) or ‘biological emotions’ (Matsumoto and Hwang 2012). These emotions are biological because they are elicited along with physiological responses from the central and autonomic nervous systems. They are psychological because they involve specific mental processes required for the elicitation and regulation of response and direct mental activities, and they

incentivize behaviour. They are social because they are often elicited by social factors and have socially communicative meaning when triggered (Matsumoto and Hwang 2012).

Emotions are rapid information processing systems that help us act with minimal conscious deliberation (Tooby and Cosmides 2008). Issues related to survival such as birth, battle, death and seduction have been present throughout our evolutionary history, and emotions aided in adapting to problems that arose rapidly with minimal conscious cognitive intervention. If we did not have emotions, we could not make rapid decisions regarding whether to attack, defend, flee, care for others, reject food or approach something useful. Think about a situation in which your friends have to rapidly avoid a car that you see is coming at them; your immediate behavioural reaction such as a fearful face or gripping your seat would quickly signal to your friends with the information that something's wrong. That immediate reaction may be more efficient than describing the danger verbally. Happiness or joy is another core emotion that is crucial for general well-being and social relationships. A smile of happiness often reduces the tension in social interactions with strangers, and happiness is the easiest way for people to express their satisfaction. Think about when you cooked your first meal for a friend and were anxious about its taste. At the moment you see your friend's smile of enjoyment, you need not even ask your friend's feedback or worry about it. Smiles often signal no opposition. Are you likely to approach your neighbour who never smiles at you? Probably not, because you instinctively know you have no signal of being accepted or affiliated with him or her.

6.1.2 The Primary Theory behind Biological Emotion

Early research on emotion was rooted in Darwin's evolutionary theory. According to Dewsbury (2009), Darwin made three main contributions to our understanding of evolution. First, Darwin offered a viable mechanism for evolutionary change and natural selection. He observed, for example, that individual variation in heritable appearances and populations are capable of increasing exponentially. Second, he claimed that evolutionary change by natural causes is totally explicable without any inherent purpose or predetermined direction. This naturalistic theory would affect many developing disciplines. Third, he accumulated massive amounts of supporting evidence across phylogenetic lines (e.g., data on beetles, barnacles, people from different cultures).

One of Darwin's interesting ideas was the continuity between human and non-human animals. Humans are different from all other species in many respects, but we also share a common origin and core, fundamental parts from some non-human primates. Much of Darwin's work was about instinctive behaviour ranging from the supernatural and theological to the natural. Darwin introduced the study of individual differences within species because such differences were basically related to natural selection.

Darwin believed that emotions and their expressions were functionally adaptive and biologically innate. In particular, Darwin ([1872] 1998) claimed in his principle of serviceable associated habits that facial expressions are the residual actions of more complete behavioural responses. For example, Darwin suggested that humans express anger by furrowing the brow and tightening the lips with teeth displayed because these actions are part of an attack response. He also believed that humans, regardless of race or culture, possessed the ability to express emotions on their faces in similar ways. Darwin engaged in a detailed study of the movements involved in emotion and concluded that those movements are universal and their antecedents can be seen in the expressive behaviours of non-human primates and other mammals. This analysis set the stage for the development of coding systems used in the identification of facial expression that have been central to the empirical literature.

In addition to the principle of serviceable associated habits, Darwin proposed an inhibition hypothesis, which stated that people are unable to perfectly simulate facial expressions in the absence of genuine emotion, and are unable to completely suppress their true expressions when feeling strong emotions, which results in emotional leakage on the face (see Baker, Black and Porter, this volume, for an argument in favour of this view). His principles and hypotheses on emotions were empirically tested and confirmed in subsequent research and lay the foundations for what are known as basic emotions.

6.1.3 *Facial Expressions of Emotion*

Of the various ways of delivering emotions, the face is one of the most complex signal systems in the body (Ekman 1972; Ekman and Friesen 1978). Our faces contain many muscles that can produce literally thousands of different types of expressions. Muscle contractions of the face are under the neural control of two different areas of the brain, one

controlling voluntary movements and the other involuntary reactions. When emotions are elicited, spontaneously produced expressions are under the neural control of the subcortical areas of the brain. When individuals attempt to control their expressions, however, those impulses are likely to stem from the cortical motor strip. Within the vast possible constellations of facial behaviours that can be produced, research has documented the existence of specific facial configurations that express emotional states (see review in Matsumoto and Juang 2013). Some facial behaviours are also signs or signals of other mental states such as cognition as well, and should also be called *expressions*. In this chapter, we explore facial expressions of emotion.

6.1.4 *Research in Facial Expressions of Emotion: Judgement Studies*

Perception of appropriate emotional expressions is crucial in social interaction (Van Kleef et al. 2011) because it often determines how to react in the situation, as well as the degree of urgency. There is no doubt about the importance of judging others' emotions and many studies have explored this topic. Tomkins conducted the first study demonstrating that facial expressions were reliably judged to be associated with certain emotional states (Tomkins and McCarter 1964), and later studies showed consistent and similar findings (Ekman, Sorenson and Friesen 1969; Izard 1971). Those initial findings were criticized, however, because the evidence for universality (i.e., the high levels of cross-cultural agreements in judgements) might have occurred because of influences of mass media (e.g., TV) and shared visual input. To address these potential limitations, Ekman and colleagues conducted two studies with two visually isolated, preliterate tribes in the highlands of New Guinea (Ekman and Friesen, 1971; Ekman, Sorenson and Friesen 1969). In their first study, the tribespeople reliably recognized facial expressions of emotion (anger, disgust, fear, happiness, surprise, sadness) posed by Westerners; in the second study, films of the tribespeople expressing emotions were shown to Americans who had never seen New Guineans before, and the Americans recognized the expressions of the New Guineans. Thus the ability to recognize facial expressions of emotion in the earlier studies did not occur because of learning through mass media or other shared visual input as the New Guineans had had no exposure to the outside world at that time.

Many subsequent studies on the universality of facial expressions of emotion examined judgements of the same facial expressions of

emotion used in the initial studies and have successfully replicated the universal recognition of emotion in the face (Matsumoto 2001; Matsumoto, Willingham and Olide 2009). Later, contempt was also identified as a universally recognized expression in various studies (Ekman and Heider 1988; Matsumoto 1992, 2005). For example, Biehl and colleagues (1997) used stimuli that validly and reliably portrayed emotional expressions (the Japanese and Caucasian Facial Expressions of Emotion – JACFEE – Matsumoto and Ekman 1988). They demonstrated that the seven emotions were reliably identified in Hungary, Japan, Poland, Sumatra, the United States and Vietnam. In addition, a meta-analysis of 168 datasets examining judgements of those emotions in the face and other nonverbal behaviours indicated universal emotion recognition well above chance levels (Elfenbein and Ambady 2002). Even when low-intensity expressions are judged, there is strong agreement across cultures in judgement (Matsumoto et al. 2002). Some of these studies included cross-cultural judgement studies of spontaneous expressions of emotions (anger, disgust, fear, disgust, fear, happiness, sadness, surprise) instead of posed ones (Matsumoto, Willingham and Olide 2009), which were convincing because of the ecological validity of the expressions tested (see Figure 6.1).

Not only are the seven universal facial expressions reliably recognized (Elfenbein and Ambady 2002; Matsumoto 2001), but cultures are similar in other aspects of emotion judgement as well. For example, there is pancultural similarity in judgements of relative intensity among faces, that is, when comparing expressions, people of different countries agree on which is more strongly expressed (Matsumoto and Ekman 1989). There is also cross-cultural agreement in the association between perceived expression intensity and inferences about subjective experiences (Matsumoto, Kasri and Kookan 1999), and in the secondary emotions portrayed in an expression (Biehl et al. 1997; Matsumoto and Ekman 1989).

6.1.5 Research in Facial Expressions of Emotion: Production Studies

Studies eliciting actual expressions of emotions on the face have been challenging to explore as doing so requires various factors and conditions in order to create the targeted emotional contexts. A perfect experimental context that produces emotions reliably may not exist. Despite that, there have been a substantial number of studies that have demonstrated that the facial expressions of anger, contempt, disgust, fear,

The Seven Universal Facial Expressions of Emotion

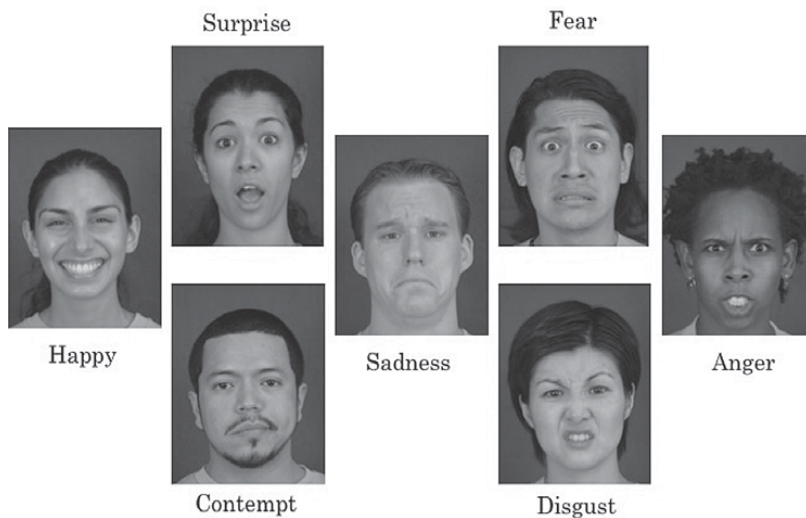


Figure 6.1. The Seven Universal Facial expressions of emotion.
Reprinted with permission © David Matsumoto.

happiness, sadness and surprise are universally produced when emotions are elicited spontaneously (see review in Hwang and Matsumoto forthcoming; Matsumoto et al. 2008).

One of the important original findings related to the universality of facial expressions of emotions came from Friesen's (1972) cross-cultural study of expressions that occurred spontaneously in reaction to emotion-eliciting films. In that study, twenty-five American and Japanese participants watched neutral and highly stressful films (first episode: body mutilation; second episode: sinus surgery) while their faces were recorded. Coding of the facial behaviours that occurred indicated that the American and Japanese participants displayed emotions (e.g., disgust, fear, sadness, anger) similarly when they were alone, and that these expressions corresponded to the facial expressions portrayed in the stimuli used in the previous judgement studies. This cross-cultural study provided pioneering evidence that facial expressions of emotion were universally produced when emotions were spontaneously elicited.

A more recent and representative production study was conducted in a natural context, which were Olympic matches. Matsumoto and

Willingham (2006) examined the expressions of eighty-four judo athletes from thirty-five countries at the 2004 Athens Olympic Games, and reported that the first immediate emotional reactions on faces of winners and losers at the completion of their final medal match were consistent with the prototypical expressions of basic emotions (see Figure 6.1). In particular, winners displayed Duchenne/genuine smiles while losers displayed sadness, disgust, anger and other negative emotions. Duchenne smiles are smiles that involve not only the smiling muscle (zygomatic major), which raises the lip corners, but also the muscles surrounding the eyes (orbicularis oculi), which raise the cheeks, thin the eyes and narrow the eye cover fold; many studies have shown that only these types of smiles are correlated with experienced positive emotion (see Ekman, Davidson and Friesen 1990 for a review).

Although there were many other cross-cultural studies of facial expressions of emotion (see methodological review in Matsumoto and Willingham [2006], which includes alternative viewpoints), this Olympic study was persuasive for several reasons; the results were based on spontaneous, not posed, expressions. Many previous studies often had to deal with variations in the basic emotions when testing posed expressions rather than using those naturally produced in an emotional context. Also, they involved individuals from various countries who participated at the Olympic games. Olympic matches can be considered as special, but they are part of our reality and are non-experimental. In that respect, the findings were meaningful given that previous research findings documenting the universality of facial expressions of emotions had been mainly tested in laboratory, not field, conditions (albeit in different laboratories using different methodologies with participants from many different cultures around the world, but all converging on the same pattern of results). Thus today there is robust evidence for the universality of seven facial expressions of emotions.

6.1.6 Evidence from Blind versus Sighted Individuals

One of the critical challenges to the notion of the biological innateness of emotion is that humans can learn and imitate emotional expressions by observing others. Blind individuals who are limited in observing and imitating others' behaviours compared to sighted people are an exceptional group in which to explore the pure effect of biologically

wired systems on the universality of emotions. This is especially true for studies involving congenitally blind individuals because they are expected to have limited social learning about how to produce sophisticated facial muscle movements of each emotion because they could not visually learn them from birth.

Previous studies have reported on the similarities between blind and sighted individuals in their spontaneous expressions of basic emotions on the face. For example, researchers have measured spontaneous facial behaviours when emotions were aroused in blind children (Cole et al. 1989). Galati, Miceli and Sini (2001) documented that there were no significant differences in displaying emotions on the face among congenitally blind and sighted young children. Later they (Galati et al. 2003) also reported that the emotional facial expressions of ten congenitally blind and ten sighted children (8–11 years old) were similar; the frequency of certain facial movements was higher in the blind children than in the sighted children.

The studies reviewed, however, did not directly compare the expressions of blind individuals from different cultures, and most tested posed, not spontaneous, facial expressions in experimental contexts rather than naturally, emotionally aroused contexts. These gaps were addressed by a later study of the spontaneous expressions of blind athletes at the 2004 Athens Paralympic Games. Matsumoto and Willingham (2009) examined spontaneous facial expressions of emotions of congenitally and non-congenitally blind judo athletes at these Games, which was a highly intense and emotional event for any athlete. The blind athletes, who came from twenty-three cultures, produced the same facial configurations of emotion documented in sighted athletes in the same emotionally evocative situations occurring in the 2004 Athens Olympic Games (see Matsumoto and Willingham 2006; see [Figure 6.2](#) for a comparative example of blind and sighted athletes' expressions).

Collectively, evidence for the similarity of facial expressions of emotions between blind and sighted individuals has been compelling and suggested that facial expressions of emotion are biologically innate. Studies of congenitally blind individuals have been especially important because it is impossible for congenitally blind individuals to produce by imitation the complicated facial expressions involved in complex muscle combinations activated spontaneously in less than a second when they experience an emotion; they would not have these automatic reactions unless they were born with the capability of experiencing and expressing the emotions in the same way as sighted

Comparison of Blind and Sighted Athletes who Just Lost a Match for a Medal

Blind athlete



Sighted athlete



(c) David Matsumoto 2009

Figure 6.2. Comparison of blind and sighted athletes who just lost a match for a medal.

Reprinted with permission © David Matsumoto.

individuals. We believe that these studies provide convincing evidence for a biologically based linkage between emotion and its expression that is universal.

6.1.7 Summary

There are many other sources of evidence that converge on the universality of facial expressions of basic emotions (readers are referred to Hwang and Matsumoto forthcoming, for a fuller discussion). According to our approach driven by Darwin's emotion theory, basic or biological emotions such as anger, contempt, disgust, fear, happiness, sadness, surprise and their expressions have evolved across species. They are evolutionarily adaptive, biologically innate and universal across humans and even non-human primates. They are primarily expressed through our faces (as well as other channels, such as voice). Much empirical evidence has supported the idea of Darwin as we discussed earlier.

6.2 Functions of Emotional Expressions on Face

In this section, we explain the roles of emotional expressions on the face as another way of understanding emotions in our lives. Although our faces deliver many emotional signals that are used for diverse functions (signalling our emotional states, illustrating or animating speech, regulating conversation, signalling thinking and concentration etc.), we focus on two main functions: intrapersonal and interpersonal effects.

6.2.1 *Intrapersonal Effects*

One of the functions that emotional expressions serve is intrapersonal, which Darwin ([1872] 1988) referred to as ‘serviceable associated habits’. For example, the raising of the brows and upper eyelid in expressions of *fear* and *surprise* allows for a widening of the visual field, allowing for individuals to see more objects in the environment. It also allows for faster eye movements during object localization, and an increase in nasal volume and air velocity during inspiration (Susskind et al. 2008). The eye widening of fear or surprise also enhances stimulus detection, whereas eye narrowing enhances stimulus discrimination (Lee et al. 2014). The dropping or opening of the jaw also allows for increased oxygen intake, priming the body for action. The wrinkling of the nose that occurs in expressions of disgust helps to close off nasal volume, allowing for a restriction of the intake of objects nasally. This makes sense if disgust is an emotion in response to contaminated objects; the closing off of the nasal passages makes the passage of contaminated objects into the body more difficult. The raising of the inner corners of the brows and muscle around the eyes in sadness stimulates the lacrimal glands, which produce tears. The lowering of the brows and glaring of the eyes in anger aids in focusing on a target, while the baring of the teeth prepares for attack responses, such as biting, a vestige of our evolutionary history.

6.2.2 *Interpersonal Effects*

Understanding others’ facial expressions of emotions often gives specific information about expressers’ emotional states, intentions, motivations and so on. Understanding biological (or basic) emotions can

be beneficial as they can be used as a universal language in social interactions regardless of different countries, cultures and languages. Emotion expressions convey important social information, for example, danger via a fearful appearance on the face, or opportunity (agreement) through a happy face (Overbeck, Neale and Govan 2010). These interpersonal effects occur because they evoke responses from others, signal the nature of interpersonal relationships and provide incentives for desired social behaviour (Keltner 2003). These interpersonal effects of facial expressions of emotion allow individuals to solve social problems. For instance, because facial expressions of emotion are universal social signals, they contain meaning not only about the expresser's psychological state but also about that person's intent and subsequent behaviour.

In social interactions the function of emotional expressions is complex; expressed emotions function not only to express senders' emotional states, but also often affect what the perceivers (or decoders) are likely to do (Hwang and Matsumoto forthcoming; Matsumoto and Juang forthcoming). People observing fearful faces, for instance, are more likely to produce approach-related behaviours, whereas people who observe angry faces are more likely to produce avoidance-related behaviours (Marsh, Ambady and Kleck 2005). This most likely occurs because angry faces increase the effectiveness and enhance the credibility of threats (Reed, DeScioli and Pinker 2014), suggesting that angry faces evolved to enhance cues of strength (Sell, Cosmides and Tooby 2014). Even the subliminal presentation of smiles produces increases in how much beverage people pour and consume and how much they are willing to pay for it; presentation of angry faces decreases these behaviours (Winkielman, Berridge and Wilbarger 2005). Emotional displays evoke specific, complementary emotional responses from observers; for example, anger evokes fear in others (Dimberg and Ohman 1996), whereas distress evokes sympathy and aid (Eisenberg et al. 1989).

Emotional expressions in social interactions also provide information about the nature of the relationships among agents. Some of the most important and provocative set of findings in this area come from studies involving married couples (Gottman and Levenson 1992; Gottman, Levenson and Woodin 2001). In this research, married couples visited a laboratory after having not seen each other for twenty-four hours, and then engaged in intimate conversations about daily events or issues of conflict. Discrete expressions of contempt, especially by the men, and disgust, especially by the women, predicted later marital dissatisfaction

and even divorce. A recent study (Overbeck, Neale and Govan 2010) tested the function of emotion in negotiation contexts and suggested that anger is helpful for powerful negotiators, and participants reported they felt more focused and assertive claiming more value; especially when the powerful party was angry, the presence of anger in the dyad (negotiation context) promoted greater value creation.

6.2.3 *Summary*

In this section, we discussed how emotions help us survive and deal with daily issues that occur in social interactions. On the intrapersonal level, emotional expressions help us identify and understand our emotional and psychological states, especially when we are sometimes unconscious about our reactions. Our spontaneous physiological reactions make it possible to manage survival risks immediately and help prepare our behaviours to react. On the interpersonal level, signalling emotions to interactants provides crucial information concerning the expressers' emotions, intentions and even attitude sometimes. The display of emotions, therefore, is not limited to showing what our emotions are; it also functions to initiate and develop social interaction, which is related to our basic needs such as creating a group, community or society (for more on this, see Helm, this volume). In the next section, we will discuss how culture aids the process and regulates emotional expressions in order to accomplish the goals and functions of emotional expressions.

6.3 Emotional Expressions and Culture

6.3.1 *Culture*

We define (human) culture as *a unique meaning and information system, shared by a group and transmitted across generations* (Matsumoto and Juang 2013) although there are various definitions of culture depending on researchers, just like emotions. One of the major functions of culture is to maintain social order, so culture creates rules, guidelines, values and norms concerning the regulation of emotion (Matsumoto and Juang 2013; see also Weisman, Bandes and Sorial, all this volume). In relation to emotions, cultures regulate biological emotions to calibrate what we become emotional about and adapt the reactions that occur when elicited. Unique human cognitive abilities including language, memory and abstract thinking allow cultures to elaborate on human emotions by facilitating the construction of culturally based emotions and their associated meanings.

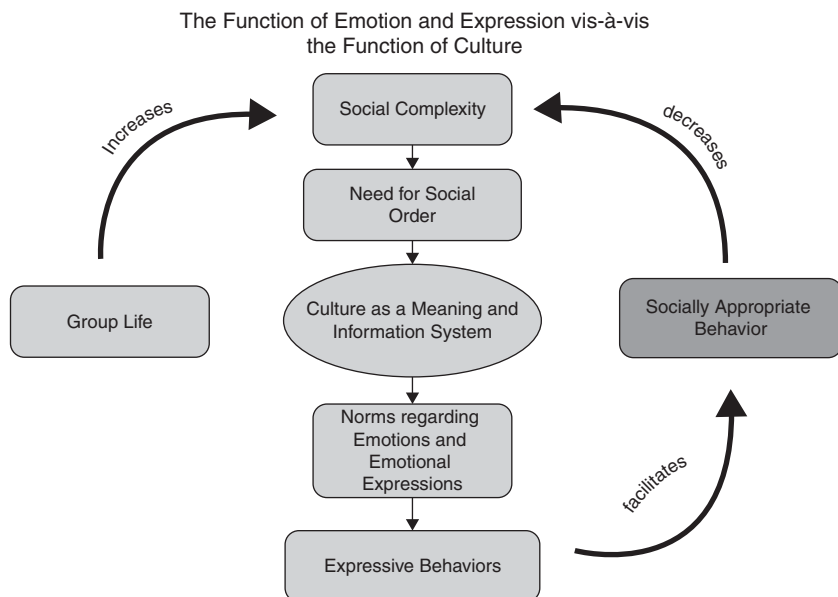


Figure 6.3. The function of emotion and expression vis-à-vis the function of culture.

Cultures create norms concerning the regulation of emotion to facilitate social coordination because emotions are primary motivators of behaviour (Tomkins 1962) and have important social functions (Keltner and Haidt 1999). Norms are guidelines for expected behaviours, thinking and feelings derived from the cultural meanings ascribed to contexts, relationships and events. They identify the range of permissible behaviour that allows groups to function effectively. Cultures encourage adherence to norms and create sanctions against infractions, which aids individuals and groups in negotiating the complexity of human social life. By regulating emotions via norms, cultures ensure that behaviours follow culturally prescribed scripts, increasing socially appropriate behaviour and thus facilitating social coordination and decreasing social chaos (see Figure 6.3; for more details see also Hwang and Matsumoto forthcoming; Matsumoto and Juang forthcoming).

6.3.2 Cultural Display Rules: Cultural Differences in Expressing Emotion

Cultural display rules are cultural norms learned early in life that govern the regulation of expressive behaviours depending on social

contexts. Emotional display rules help us manage the appropriateness of emotional displays in social situations when there is the need to manage the appearance of particular emotions in particular situations (Diefendorff and Greguras 2009; Matsumoto 1990, 1993). In relation to emotion knowledge, which refers to familiarity with social display rules, display rules are conventions that determine when, where, and how expressive behaviours should be conveyed (Saarni 1984). The rules differ across cultures, contexts and interactants.

The term 'cultural display rules' was coined by Ekman and Friesen (1969) to explain how universal facial expressions of emotion could be managed and even modified according to social context by people of different cultures. They conducted a classic study to demonstrate the existence of display rules (the first part of which was described earlier in this chapter). In that study (Friesen 1972), American and Japanese participants viewed stressful films alone and in the presence of an experimenter. When alone, they displayed the same expressions of disgust, anger, fear and sadness. When they viewed the films a second time with an experimenter, however, there were dramatic differences. While the Americans tended to continue to show their negative feelings, many Japanese smiled. The researchers concluded that cultural display rules were operating, which prevented the free expression of negative emotions in the presence of another person in the Japanese culture.

Cultural display rules are important because they explain how cultures can influence a biologically based, innate ability like universal facial expressions of emotion. There are multiple ways in which display rules can act to modify expressions: (1) Express *less* than actually felt (Deamplification); (2) Express *more* than actually felt (Amplification); (3) Show nothing (Neutralization); (4) Show the emotion but with another emotion to comment on it (Qualification); (5) Mask or conceal feelings by showing something else (Masking); (6) Show an emotion when they really don't feel it (Simulation). These rules explain how facial expressions of emotion can be both universal and culture-specific. They govern how universal emotions can be expressed and centre on the appropriateness of displaying each of the emotions in particular social circumstances. They are learned early and dictate how the universal emotional expressions should be modified according to the social situation. By adulthood, these rules are automatic, having been very well practiced.

A later study on display rules on emotions by Matsumoto (1990) was meaningful as it was the first cross-cultural study that directly

measured differences between two cultures. In that study, Americans and Japanese participants saw two examples of six universal facial expressions of emotion and rated the appropriateness of displaying each in eight social contexts. Americans rated some negative emotions in ingroups, and happiness in outgroups, more appropriate than did the Japanese. The Japanese, however, rated some negative emotions as more appropriate to outgroup members. These findings were interpreted within an individualism versus collectivism framework, reckoning that members of collectivistic cultures (i.e., Japan) would be discouraged from expressing potentially threatening negative emotions to their ingroups, and potentially bonding positive emotions to their outgroups, while there would be no such tendency for members of individualistic cultures (i.e., the United States).

In relation to individual variations in display rules, Hudson and Jacques (2014) reported that children (5–7 years old) were capable of regulating their negative emotions in front of parents in a disappointing gift paradigm. Age was a reliable predictor of overt effort to regulate emotions. This finding implied that not only are there cultural differences in display rules, but also age differences as well, as individuals learn the social rules and manners over time.

6.3.3 The Display Rule Assessment Inventory and a Worldwide Mapping of Cultural Display Rules

After the original inception of the concept of display rules, cross-cultural research on them was dormant until Matsumoto's (1990) study examining display rules in Americans and Japanese (described earlier), and a similar study documenting differences in display rules among four ethnic groups within the United States (Matsumoto 1993). Later, Matsumoto and colleagues created the Display Rule Assessment Inventory (DRAI), where participants choose one of six behavioural responses (corresponding to the ways expressions are managed in real life, as described) when they experience different emotions with family, friends, colleagues and strangers (Matsumoto et al. 1998, 2005). The behavioural responses were based on the six response alternatives described. They demonstrated cultural differences in display rules and provided evidence for the DRAI's internal and temporal reliability and for its content, convergent, discriminant, external and concurrent predictive validity with personality.

Matsumoto and colleagues (2008) then administered a more comprehensive version of the DRAI in over thirty countries, examining

universal and culture-specific aspects to display rules and linking the cultural differences to culture-level individualism (vs. collectivism). Despite the larger potential range of scores, most countries' means on overall expression endorsement suggested a universal norm for expression management. Individuals of all cultures had a display rule norm for greater expressivity towards ingroups than towards outgroups, indicating another universal effect. Collectivistic cultures were associated with a display rule norm of less expressivity overall than individualistic cultures, suggesting that overall expression management for all emotions is central to the preservation of social order in these cultures. This finding is commensurate with the behavioural findings from previous studies (Friesen 1972; Matsumoto, Willingham and Ollide 2009). Individualism was also positively associated with higher expressivity norms in general and for positive emotions in particular. Individualism was positively associated with endorsement of expressions of all emotions towards ingroups, but negatively correlated with all negative emotions and positively correlated with happiness and surprise towards outgroups.

A more recent study (Hwang and Matsumoto 2012) using the DRAI reported that Asian Americans and European Americans had different display rules for their emotional expressions. Asian Americans reported expressing their emotions less than European Americans, but modifying their expressions more. However, perceived relationship commitment ratings mediated the ethnic group variations on endorsed expressivity. European Americans had significantly higher scores on perceived relationship commitment than did Asian Americans towards parents, close friends and older siblings, and these differences completely mediated ethnic differences in display rules.

6.4 Conclusion

Humans have emotions that have evolved across species, and are adaptive, biologically innate and universal (Darwin [1872] 1998). Research has supported this idea with empirical findings as reviewed in the preceeding text. That humans share certain similarities related to the survival of the species since human history began is reasonable and entirely functional. Understanding that emotional signals on the face are applicable across cultures, ethnicities and gender is beneficial in our daily life and social interactions. However, we do not argue that all humans are exactly the same and there is no difference and variation in displaying emotions at

all; there is of course a wide range of individual differences in emotional expressivity and in how cultures regulate emotions and their expressions.

Human cultures perform an important role in moderating the universal facial signals of emotions once they are experienced and expressed. Culture is an essential guideline for people to socially modify their more voluntary display rule driven facial behaviours, which occur after immediate behavioural reactions, in order to smoothen their social interactions for social coordination. Culture more heavily influences learned expressions such as those associated with speech illustration or emblematic information. Also, the effects of context should not be neglected, because universal behavioural reactions and cultural specific management of those reactions occur in specific contexts. Understanding and identifying the ultimate and essential reason why we have emotions and signal them in certain ways to others and ourselves gives us a pathway to study emotion and its expressions in research and to improve our social communications for us and the people surrounding us. To utilize this valuable contribution of emotion in our daily life, future research will need to continue to expand the cross-cultural studies in expressions of emotions.

References

- Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T. et al. 1997. 'Matsumoto and Ekman's Japanese and Caucasian Facial Expressions of Emotion (JACFEE): Reliability Data and Cross-National Differences', *Journal of Nonverbal Behavior* 21: 3–21.
- Cole, P. M., Jenkins, P. A. and Shott, C. T. 1989. 'Spontaneous Expressive Control in Blind and Sighted Children', *Child Development* 60: 683–688.
- Darwin, C. [1872] 1998. *The Expression of Emotion in Man and Animals*. New York: Oxford University Press.
- Dewsbury, D. A. 2009. 'Charles Darwin and Psychology at the Bicentennial and Sesquicentennial: An Introduction', *American Psychology* 64: 67–74.
- Diefendorff, J. M. and Greguras, G. J. 2009. 'Contextualizing Emotional Display Rules: Examining the Roles of Targets and Discrete Emotions in Shaping Display Rule Perceptions', *Journal of Management* 35: 880–898.
- Dimberg, U. and Ohman, A. 1996. 'Behold the Wrath: Psychophysiological Responses to Facial Stimuli', *Motivation and Emotion* 20: 149–182.
- Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M. et al. 1989. 'Relation of Sympathy and Distress to Prosocial Behavior: A Multimethod Study', *Journal of Personality and Social Psychology* 57: 55–66.
- Ekman, P. 1972. 'Universal and Cultural Differences in Facial Expression of Emotion', in Cole (ed.) *Nebraska Symposium on Motivation*, 1971. Lincoln: Nebraska University Press, pp. 207–283.

- Ekman, P. and Friesen, W. V. 1969. 'The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding', *Semiotica* 1: 49–98.
1971. 'Constants across Culture in the Face and Emotion', *Journal of Personality and Social Psychology* 17: 124–129.
1978. *Facial Action Coding System: Investigator's Guide*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P. and Heider, K. G. 1988. 'The Universality of a Contempt Expression: A Replication', *Motivation and Emotion* 12: 303–308.
- Ekman, P., Davidson, R. J. and Friesen, W. V. 1990. 'The Duchenne Smile: Emotional Expression and Brain Physiology: II', *Journal of Personality and Social Psychology* 58: 342–353.
- Ekman, P., Sorenson, E. R. and Friesen, W. V. 1969. 'Pancultural Elements in Facial Displays of Emotion', *Science* 164: 86–88.
- Elfenbein, H. A. and Ambady, N. 2002. 'On the Universality and Cultural Specificity of Emotion Recognition: A Meta-Analysis', *Psychological Bulletin* 128: 205–235.
- Friesen, W. V. 1972. *Cultural Differences in Facial Expressions in a Social Situation: An Experimental Test of the Concept of Display Rules*. (Doctoral dissertation). University of California, San Francisco.
- Galati, D., Miceli, R. and Sini, B. 2001. 'Judging and Coding Facial Expression of Emotions in Congenitally Blind Children', *International Journal of Behavioral Development* 25: 268–278.
- Galati, D., Sini, B., Schmidt, S. and Tinti, C. 2003. 'Spontaneous Facial Expressions in Congenitally Blind and Sighted Children Aged 8–11', *Journal of Visual Impairment and Blindness* 97: 418–428.
- Gottman, J. M. and Levenson, R. W. 1992. 'Marital Processes Predictive of Later Dissolution: Behavior, Physiology, and Health', *Journal of Personality and Social Psychology* 63: 221–223.
- Gottman, J. M., Levenson, R. W. and Woodin, E. 2001. 'Facial Expressions during Marital Conflict', *Journal of Family Communication* 1: 37–57.
- Hudson, A. and Jacques, S. 2014. 'Put on a Happy Face! Inhibitory Control and Socioemotional Knowledge Predict Emotion Regulation in 5- to 7-Year-Olds', *Journal of Experimental Child Psychology* 123: 36–52.
- Hwang, H. S. and Matsumoto, D. 2012. 'Ethnic Differences in Display Are Mediated by Perceived Relationship Commitment', *Asian American Journal of Psychology* 3: 254–262.
- forthcoming. 'Measuring Emotions in the Face', in Meiselman (ed.) *Emotion Measurement*. Cambridge, UK: Elsevier.
- Izard, C. E. 1971. *The Face of Emotion*. East Norwalk, CT: Appleton-Century-Crofts.
- Keltner, D. 2003. 'Expression and the Course of Life: Studies of Emotion, Personality, and Psychopathology from a Social-Functional Perspective', in Ekman, Campos, Davidson and De Waal (eds.) *Emotions Inside Out: 130 Years after Darwin's 'The Expression of the Emotions in Man and Animals'*, Vol. 1000. New York: Annals of the New York Academy of Sciences, pp. 88–90.
- Keltner, D. and Haidt, J. 1999. 'Social Functions of Emotions at Four Levels of Analysis', *Cognition and Emotion* 13: 505–522.

- Lee, D. H., Mirza, R., Flanagan, J. G. and Anderson, A. K. 2014. 'Optical Origins of Opposing Facial Expression Actions', *Psychological Science* 25: 745–752.
- Marsh, A. A., Ambady, N. and Kleck, R. E. 2005. 'The Effects of Fear and Anger Facial Expressions on Approach- and Avoidance-Related Behaviors', *Emotion* 5: 119–124.
- Matsumoto, D. 1990. 'Cultural Similarities and Differences in Display Rules', *Motivation and Emotion* 14: 195–214.
1992. 'American-Japanese Cultural Differences in the Recognition of Universal Facial Expressions', *Journal of Cross-Cultural Psychology* 23: 72–84.
1993. 'Ethnic Differences in Affect Intensity, Emotion Judgments, Display Rule Attitudes, and Self-Reported Emotional Expression in an American Sample', *Motivation and Emotion* 17: 107–123.
2001. 'Culture and Emotion', in Matsumoto (ed.) *The Handbook of Culture and Psychology*. New York: Oxford University Press, pp. 171–194.
2005. 'Scalar Ratings of Contempt Expressions', *Journal of Nonverbal Behavior* 29: 91–104.
- Matsumoto, D. and Ekman, P. 1988. *Japanese and Caucasian Facial Expressions of Emotion (JACFEE) [Slides]*. Francisco, CA: Intercultural and Emotion Research Laboratory, Department of Psychology, San Francisco State University.
1989. 'American-Japanese Cultural Differences in Intensity Ratings of Facial Expressions of Emotion', *Motivation and Emotion* 13: 143–157.
- Matsumoto, D. and Hwang, H. S. 2012. 'Culture and Emotion: The Integration of Biological and Cultural Contributions', *Journal of Cross-Cultural Psychology* 43: 91–118.
- Matsumoto, D. and Juang, L. 2013. *Culture and Psychology*, 5th edn. Belmont, CA: Wadsworth/Cengage Learning.
- forthcoming. *Culture and Psychology*, 6th edn. Belmont, CA: Wadsworth/Cengage Learning.
- Matsumoto, D. and Willingham, B. 2006. 'The Thrill of Victory and the Agony of Defeat: Spontaneous Expressions of Medal Winners of the 2004 Athens Olympic Games', *Journal of Personality and Social Psychology* 91: 568–581.
2009. 'Spontaneous Facial Expressions of Emotion of Congenitally and Noncongenitally Blind Individuals', *Journal of Personality and Social Psychology* 96: 1–10.
- Matsumoto, D., Kasri, F. and Kookan, K. 1999. 'American-Japanese Cultural Differences in Judgments of Expression Intensity and Subjective Experience', *Cognition and Emotion* 13: 201–218.
- Matsumoto, D., Willingham, B. and Olide, A. 2009. 'Sequential Dynamics of Culturally Moderated Facial Expressions of Emotion', *Psychological Science* 20: 1269–1274.
- Matsumoto, D., Yoo, S. H., Hirayama, S., and Petrova, G. 2005. 'Validation of an Individual-Level Measure of Display Rules: The Display Rule Assessment Inventory (DRAI)', *Emotion* 5: 23–40.
- Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M. and Frank, M. G. 2008. 'What's in a Face? Facial Expressions as Signals of Discrete Emotions', in

- Lewis, Haviland-Jones and Barrett (eds.) *Handbook of Emotions*, 3rd edn. New York: Guilford Press, pp. 211–234.
- Matsumoto, D., Olide, A., Schug, J., Willingham, B. and Callan, M. 2009. 'Cross-Cultural Judgments of Spontaneous Facial Expressions of Emotion', *Journal of Nonverbal Behavior* 33: 213–238.
- Matsumoto, D., Takeuchi, S., Andayani, S., Kouznetsova, N. and Krupp, D. 1998. 'The Contribution of Individualism-Collectivism to Cross-National Differences in Display Rules', *Asian Journal of Social Psychology* 1: 147–165.
- Matsumoto, D., Consolacion, T., Yamada, H., Suzuki, R., Franklin, B., Paul, S. et al. 2002. 'American-Japanese Cultural Differences in Judgements of Emotional Expressions of Different Intensities', *Cognition and Emotion* 16: 721–747.
- Matsumoto, D., Yoo, S. H., Fontaine, J. R. J., Anguas-Wong, A. M., Arriola, M., Ataca, B., ... Grossi, E. 2008. 'Mapping Expressive Differences around the World: The Relationship between Emotional Display Rules and Individualism v. Collectivism', *Journal of Cross-Cultural Psychology* 39: 55–74.
- Overbeck, J., Neale, M. and Govan, C. 2010. 'I Feel, Therefore You Act: Intrapersonal and Interpersonal Effects of Emotion on Negotiation as a Function of Social Power', *Organizational Behavior and Human Decision Processes* 112: 126–139.
- Reed, L. I., DeScioli, P. and Pinker, S. A. 2014. 'The Commitment Function of Angry Facial Expressions', *Psychological Science* 25: 1511–1517.
- Saarni, C. 1984. 'An Observational Study of Children's Attempts to Monitor Their Expressive Behavior', *Child Development* 55: 1504–1513.
- Sell, A., Cosmides, L. and Tooby, J. 2014. 'The Human Anger Face Evolved to Enhance Cues of Strength', *Evolution and Human Behavior* 35: 425–429.
- Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W. and Anderson, A. K. 2008. 'Expressing Fear Enhances Sensory Acquisition', *Nature Neuroscience* 11: 843–850.
- Tomkins, S. S. 1962. *Affect, Imagery, and Consciousness, Vol. 1: The Positive Affects*. New York: Springer.
- Tomkins, S. S. and McCarter, R. 1964. 'What and Where Are the Primary Affects? Some Evidence for a Theory', *Perceptual and Motor Skills* 18: 119–158.
- Tooby, J. and Cosmides, L. 2008. 'The Evolutionary Psychology of Emotions and Their Relationship to Internal Regulatory Variables', in Lewis and Haviland-Jones (eds.) *Handbook of Emotions*, 3rd edn. New York: Guilford.
- Van Kleef, G., Van Doorn, E., Heerdink, M. and Koning, L. 2011. 'Emotion Is for Influence', *European Review of Social Psychology* 22: 114–163.
- Winkielman, P., Berridge, K. C. and Wilbarger, J. L. 2005. 'Unconscious Affective Reactions to Masked Happy versus Angry Faces Influence Consumption Behavior and Judgments of Value', *Personality and Social Psychology Bulletin* 31: 121–135.

7. A Sceptical Look at Faces as Emotion Signals

James A. Russell

As the authors of other chapters in this volume elaborate, it is common sense that we express our emotions in our face and that onlookers can read the emotions so expressed. These seemingly obvious premises stem from a folk theory that dates back at least to Aristotle. They were articulated into a simple and heuristic scientific research program known as Basic Emotion Theory (BET) and that has dominated the psychological study of emotion for decades. The theory was readily accepted both within and outside science. Researchers, clinicians and security guards accepted its advice on how to detect emotions by watching for tell-tale facial signs. Philosophers relied on its claims in analyzing emotion and expression (see, e.g., Green 2007 and this volume). Alas, as often happens in science, the evidence gathered indicates that BET must be revised or, as I suggest, abandoned altogether. I first review that evidence and then turn, briefly, to consider a way to think about facial changes that does not assume that emotions cause them.

7.1 The Prototype Version of Basic Emotion Theory

In the early decades of the twentieth century, experimental psychologists sought to bring into the laboratory the age-old idea that all human beings express their emotions by the same facial movements and that onlookers can accurately read those emotions from those expressions. Indeed, Langfeld (1918a: 183) called this idea ‘a well-known fact’. The experimentalists were therefore presumably surprised when the evidence was mixed. A person’s emotion did not reliably result in a specific facial movement (Landis 1924), and onlookers were often inaccurate in reading emotions even from highly selected facial expressions (Buzby 1924; Feleky 1914; Guilford 1929; Langfeld 1918a, 1918b;

Sherman 1927). Anthropologists observed cultural differences in how emotions were expressed (Bateson and Mead 1942).

In 1962, Tomkins published the first of four volumes that proposed what I shall call BET. BET was a return to the millennia-old common-sense folk theory of human facial expressions, although BET enthusiasts sometimes date this theory to a much more recent event, the publication of Charles Darwin's (1872) book on the expression of emotion. BET captures our taken-for-granted presuppositions about facial expressions. BET articulated one version of the folk theory and added an evolutionary history, a neural mechanism (called an 'affect program', which is a computer metaphor for a hard-wired circuit in the brain), and supported by a famous set of studies conducted in the highlands of Papua New Guinea (for discussion of these experiments, see Hwang and Matsumoto, this volume). The evolutionary account explained emotions as pre-prepared solutions to recurring problems encountered during the Era of Evolutionary Adaptedness. Each pre-prepared solution (each basic emotion) included a conscious feeling, an emotion-specific pattern of autonomic nervous system changes, a facial signal and an overt behavioural response, all cohering in a tightly coordinated package. Led by the empirical work of Ekman (1972) and Izard (1971), BET came to be the account presented in the textbooks and the dominant research program on emotion in psychology.

Scientific theories often emerge out of a folk theory, but then change, sometimes in fundamental ways, as research uncovers unpredicted facts and anomalies. Casual observation, and hence folk theory, suggested to early scientists a flat earth under the sky. The sky appears to be a single sphere in which reside the sun, moon, planets and stars. Analysis of ordinary language phrases such as 'sunrise' and 'sunset' would have supported this theory. Early astronomers made more careful observations and hypothesized eight different spheres (one for the moon, another for the sun and so on), which allowed better predictions of such events as eclipses. Improved predictions required more spheres of different shapes. The process of revising the geocentric theory became more and more complicated, until of course Copernicus, Galileo and Newton.

Similarly, casual observations suggest that we smile when happy, frown when angry and so on. Early experimental research suggested modifications to the initial theory (such as display rules and microexpressions – see Hwang and Matsumoto, this volume, for discussion of the former and Baker et al., this volume, for discussion of

the latter). Observations and experiments continued to uncover unpredicted facts and anomalies about faces that demand a similar move away from folk theory and this prototypical version of BET. I next summarize that evidence, first, concerning the production of facial expressions and, second, concerning the recognition of emotion from them.

7.2 The Production of Facial Movements

All people – whatever their age, education or culture – move their faces: we smile, frown, grimace, wince, gasp and so on. But, that fact alone does not tell us what causes us to do so or the internal mechanisms involved. The first task is to account for various facial movements. Common sense provides an answer for some movements: happiness makes us smile, anger makes us frown, fear makes us gasp, disgust makes us scrunch our noses and so on. BET formulated these common-sense assumptions as scientific hypotheses, but has offered surprisingly little direct evidence to support those hypotheses. And, indeed, the available evidence is not encouraging to BET (although, for a contrary view, see Hwang and Matsumoto, this volume).

Reisenzein, Studtmann and Horstmann (2013) reviewed the laboratory evidence and Fernández Dols and Crivelli (2013) the field evidence. The most readily agreed upon association is that between happiness and smiles, and yet the association is much weaker than BET implied: happy people not smiling, and smiles without happiness. Smiles are easily posed, do not correlate well with the smiler's emotional state (Krumhuber and Manstead 2009) and can be caused by negative experiences such as losing a game (Schneider and Josephs 1991), embarrassment (Keltner and Anderson 2000) or pain (Kunz, Prkachin and Lautenbacher 2009). In the most important study of happiness and smiling, Fernández Dols and Ruiz-Belda assessed smiling in gold medal winners at Olympic games. Although winners rated themselves as extremely happy, they smiled during specific social events but not at other times. Similar problems arise for other emotion–face associations. For example, Reisenzein (2000) created various laboratory situations that elicited surprise and observed the facial behaviour of the surprised participants. BET's predicted 'surprise expression' (raised brow and widened eyes) occurred but rarely.

Science needs to begin with a broader inventory of the plausible causes and mechanisms of facial movement. Our faces move as we eat, breathe, exert effort, smell, feel pain or reach orgasm. Our faces move

during certain reflexes (gag, orienting, startle etc.) and during perception (looking, tasting etc.). Our faces are especially active during social interaction: our faces move as we talk and gesture (social greeting, threatening, exerting dominance or submission). Our faces move as we unconsciously imitate others and as we deliberately negotiate interaction with others. Our faces move according to cultural prescriptions and proscriptions (faking, masking, hiding; in my culture, downcast eyes at a funeral; smiles when greeting someone); these cultural rules were hypothesized by Klineberg (1938) to account for cultural differences in facial behaviour and later adopted by Ekman (1972) under the name 'display rules'. (Whether the rules are in fact determined by nurture rather than nature is an open question; Griffiths 2003.)

Such an inventory is a first step in a functional analysis, but it is also important in an empirical assessment of BET. The question is whether various non-emotional sources can cumulatively account for facial movement attributed by BET to emotion. Non-emotional processes provide possible alternatives to BET's emotion account of facial movements. The alternative accounts would be of this form: the 'facial expression of surprise' includes raised eyebrows and widened eyes, which aid in visual scanning, but visual scanning alone even without surprise includes raising the eyebrows and widening the eyes; surprise without visual scanning would not. Tasting and smelling can wrinkle the nose and raise the upper lip – whether or not one is disgusted – creating the muscle movements alleged by BET to signal disgust. Showing a face that BET might call shame or guilt need not be accompanied by shame or guilt but might occur as a strategy for repairing an important relationship. The rapid facial movement seen during a social interaction points to a neglected phenomenon not plausibly attributed to shifting emotions.

Appraisal theories of emotion can provide a conceptual framework in which facial movement is correlated with but not caused by emotion (Ortony and Turner 1990). A perceptual-cognitive process includes facial movements and also leads to the experience of emotion. For example, perceiving the current situation as puzzling increases attention and furrows the brow and, sometimes, leads to the experience of surprise. The facial movement is correlated with an emotion but is caused by a perceptual rather than emotional process. Similarly, a smile might be correlated with happiness but caused by a social interaction. To test this theory requires us to ask what happens when the social interaction and the emotion are separated. Kraut and Johnston's (1979)

study of smiling bowlers took up this question. Kraut and Johnston agreed that people often smile when happily interacting with others, but asked what happens when happiness and interacting are teased apart. Smiles when happy but not interacting were found to be rare. No one study is definitive, but follow-up studies with changes in method yielded similar results (Crivelli, Carrera and Fernández-Dols 2015; Fernández Dols and Ruiz-Belda 1995; Fernández-Dols, Carrera and Crivelli 2011; Ruiz-Belda, Fernández Dols, Carrera and Barchard 2003). There are also theories of facial movement that omit the concept of emotion. The most promising alternative to an emotion account of facial movement is based in modern evolutionary theory (Fridlund 1994; see Parkinson 2005, for a review). The general idea is that facial movements commonly thought to be expressions of emotion occur as negotiation between individuals. Rather than an emotion, the cause of the facial movement is a behavioural intention (or threat). A smile, for example, is a signal to a specific audience of a readiness to interact. A 'sad face' is a signal requesting comfort or aid. An 'anger frown' is a signal threatening hostility and might well be a bluff. Fridlund's theory abandoned the implication of altruistic signalling implied by BET and emphasized the conflict of interests between signaller and audience. Fridlund's theory is especially important in developing an account of the rapidly changing facial movements that guide and respond during a quickly unfolding social interaction.

One counter-argument to Fridlund's theory was that we smile at pleasant thoughts even when alone. Fridlund (1994) showed that even when alone we nonetheless smile at an audience – an imaginary one – and smiling declines as the implicit audience becomes less salient. (We would also need to study actual cases of being alone, rather than fictional cases created with intuitions or memories.) Another counter-argument was that behavioural intention is part of rather than an alternative to emotion. On BET, however, a specific behaviour was pre-programmed during the Era of Evolutionary Adaptedness, and the signal is broadcast to all who happen to see it. In contrast, for Fridlund the behavioural intention is currently constructed as part of negotiation with a specific audience, and sending the signal is not altruistic. Initially, actors may attempt to manipulate audiences with their displays of intent (which may be overplayed as 'bluffs' or underplayed as enticements). What displays occur depend not just on the current situation but on the history of interaction and can change rapidly. Fridlund's theory thus emphasizes that facial signalling is thoroughly dynamic,

embedded within an ongoing social context, responsive to the interaction partner and dependent on history with that audience and anticipation of the future. Fridlund's concept of behavioural intention thus differs considerably from BET's notion of what behaviour is contained in a basic emotion.

In short, because common sense assumes that emotions cause facial expressions, we often think we see emotions in the facial movements we witness. That interpretation seems intuitive, even obvious. (Fictional examples that engage these intuitions are easily concocted.) Yet scientific evidence is lacking to substantiate this interpretation. Although BET has been the dominant scientific programme for emotion for five decades, it has offered no convincing evidence that emotions *cause* facial movements. Emotion is sometimes correlated with its predicted facial expression, but the correlation has been found to be weak, and emotion is confounded with other sources. The existence of other sources for smiles does not rule out that happiness might sometimes cause smiles, but, however, we have no evidence that it does. When experiments disentangle emotion from those confounding sources, emotion does not appear to be the causal agent. Therefore, we must seek alternative explanations to the claim that an emotion is causing the face to move.

7.3 The Onlooker's Interpretation of Facial Movements

Evidence offered for BET focuses less on the production of facial expressions than on the alleged 'recognition of emotion' from them. BET articulated the common belief that people 'recognize' happiness, anger, disgust and so on in the faces of others. The specific emotion that is conveyed by BET's predicted facial expressions has been characterized as 'obvious' (Ekman 1980: 11), and recognition of those emotions as 'universal' (Ekman and Cardaro 2011), 'automatic' (Tracy and Robins 2008: 81), 'easy' (Izard 1971: 502) and even 'innate' (Martens, Hamlin and Tracy 2013). Here I argue that the evidence does not support BET's account of recognition of emotion.

Recognition is a loaded word in that it implies that the happiness, anger and so on are there to be recognized. But even if universal consensual agreement could be established, doing so would not establish the validity of what is agreed upon: we may consensually think that we see the sun rise and set and think that we hear joy in birdsong; but, of course, the sun does not actually rise or set, and birds do not sing out of joy.

Many studies purported to demonstrate the predicted 'recognition' by asking people to match a photo of a posed facial expression to one of BET's predicted emotion terms (see Nelson and Russell 2013; Russell 1994, for reviews). Such demonstrations, even if reliable, would not show that people's emotion attributions are accurate. Nor would they show that people spontaneously perceive the predicted emotion but that, once told that one of a small number of emotions is expressed, they can select the predicted one. Even more troubling, the high 'recognition' (matching) scores found may be partly due to design methods. No one design factor need be powerful (although they often are), but cumulatively they combine to push scores toward the predicted outcome. In the typical study, observers see many faces one at a time (a within-subjects experimental design); when observers instead see just one face, matching scores are lower. In the typical study, observers are shown posed exaggerated facial expressions; when observers see spontaneous rather than posed faces, matching scores plummet. In the typical study, the observer must select one and only one option from a multiple-choice list (forced-choice response format); when observers are allowed to choose their own label, matching scores drop. Indeed, with a forced-choice response format, when the list is changed, the consensual label can change. Design factors can even be powerful enough to create false recognition. We recently found evidence that people can achieve a high 'recognition score' on a nonsense label for a facial expression: both children and adults recognized 'pax', a made-up word, as the emotion expressed by a face (DiGirolamo and Russell 2014; Nelson and Russell 2014). Observers initially were shown prototypical facial expressions of basic emotions, and they matched faces to BET's predicted emotion. When they next saw a new facial expression, they tended to choose the word from the response format that they had not previously used – even when the word was nonsense. Our interpretation was that they used a process of elimination to select the word for the face. If so, consensual recognition of emotion from some (but not all) faces could similarly have been aided by a process of elimination.

The studies commonly cited as support for BET's prediction lack ecological validity. In the typical study (e.g., Ekman and Friesen 1971), the observer is shown a set of photos selected by the experimenter. Although the photos are said to represent facial expressions, each expression exists on a face with a certain physiognomy. Thus, expression and physiognomy are typically confounded in the stimulus. The observer sees a static photo of the face at an optimal distance for visual

perception, and the face is devoid of any context (the expresser's body, voice, movements and external situation). The observer in a real-life situation almost never uses someone's facial expression alone and at a slice in time to infer that person's emotion, but facial expression as it changes against its physiognomic background in response to its context and accompanied by whatever else the expresser is doing (see Hess and Hareli, this volume, for a discussion of the role of context in the interpretation of facial expressions). So, removing motion and context stands in the way of understanding how observers interpret facial expressions in the non-laboratory world.

More important, context is more powerful for the recognition of emotion than BET predicts. According to BET, the facial expression is automatically produced when the emotion is activated and is automatically recognized. This characterization seems to imply that the facial expression trumps context, which can provide only probabilistic information because different individuals respond differently to the same situation. In fact, context trumps facial expression (Carroll and Russell 1996): a person in an anger-inducing situation but showing BET's 'fear face' was interpreted as angry rather than afraid. Contrary to my own prediction, context trumps face even on valence (whether the emotion is pleasant or unpleasant; Aviezer et al. 2008; Kayyal, Widen and Russell 2015).

BET emphasizes the universality of recognition: basic emotions were claimed to be recognized from the predicted facial expressions by all people whatever their culture, language or education. Yet, a meta-analysis found that recognition (matching) scores vary with culture, language and education (Nelson and Russell 2013; Trauffer, Widen and Russell 2013). Jack et al. (2012) used a psychophysical technique and again found cultural differences. BET presupposed that the English words *happiness*, *anger*, *disgust* and so on express universal categories in terms of which recognition proceeds; evidence challenges this assumption (Russell 1991; Wierzbicka 1999): different languages do not have terms that exactly match the English terms, although there are similarities. As Ekman and Friesen (1971) emphasized, the most telling test of universality involves societies remote from Western culture and media. The few studies of remote societies showed a large cultural difference (Nelson and Russell 2013; Russell 1994). Recently, Gendron et al. (2014) similarly found weak to non-existent support for BET's predictions in a remote society. I believe that people everywhere have similar interpretations of facial movements (a position I termed 'minimal universality', Russell 1995), but not in the way predicted by BET.

Some writers emphasize that BET's hypotheses are supported to a statistically significant degree (e.g., Shariff and Tracy 2011). For example, observers select BET's predicted emotion label more often than they would if they chose labels randomly. (Whether the same can be said when method artefacts are removed is unknown.) No one predicts that humans are random in interpreting faces. Ruling out the null hypothesis of random responding does not rule in the experimenter's hypothesis. There are many ways to explain non-random responding; I offered eight alternative accounts, and surely there are more (Russell 1994), and all of them predict non-random responding. (That the geocentric theory of the solar system can make predictions about the time of sunrise that are much more accurate than chance does not prove the geocentric theory.)

The process called 'recognition' is interpretation. Observers use facial information to make inferences not just about emotions but about any psychological state. The interpretation of the face is influenced by various factors, some rarely studied (e.g., colour of the sclera), some more studied: by the observer's situation (state, interests, motives), by the face's context (the expresser's context, words, gaze, vocal prosody, body position and proxemics, motor behaviour and underlying physiognomy) and by features of the experimental method. Even with regard to emotion, we need to study not six English folk terms for emotion (*happiness, anger, disgust* etc.) but the many categories of emotion and how their accessibility or even existence varies with language and culture.

In short, sufficient evidence has now accumulated to know that BET's claims about universal recognition of emotion from facial expression are unwarranted. I have no *a priori* argument that BET cannot be a valid account of recognition of emotion from facial expressions, but its prognosis on the available evidence is poor.

7.4 Revisions to BET

One response to the evidence mentioned here might be to revise BET. This tack must be evaluated in light of evidence on other aspects of the theory. There is no consensually agreed upon confirmatory evidence for emotion-specific signatures in the autonomic nervous system (Cacioppo et al. 2000) or specific behavioural responses (Baumeister et al. 2007). BET predicts tight coherence among each emotion's components, but they turn out to be surprisingly weakly correlated

(Reisenzein 2000). Modern understanding of evolution through natural selection raises doubts about BET and provides an alternative (Buss 2014; Fridlund 1994). Automatic signalling of one's true emotion to enemies could incur heavy costs, and evolution is likely to have produced deceptive as well as veridical signals because of conflict of interest between expresser and observer.

Yet another caution in revising BET is that revisions may introduce problems as much as solutions. For example, cultural differences led Ekman (1972) to embrace Klineberg's (1938) hypothesis of cultural rules prescribing or proscribing facial expressions. There are such cultural rules, and people attempt to conform to them. On Ekman's (1972) treatment, however, these display rules render his theory immune to evidence. Happiness leads to smiles or at least micro-smiles, except when it doesn't, in which case a display rule intervened. Without a prior specification of the display rules, no evidence could falsify the theory. Scarantino (2015) proposed a New BET, but with new categories of emotion that remain to be specified.

7.5 The Nature of Emotion

Attempts to revise BET face a deeper problem. BET's empirical difficulties expose problems with the background assumptions on which it was built. BET borrowed the everyday concept of emotion, but it is not clear just what an emotion is and whether it is the sort of thing that can cause facial expressions. In the previous sections, I wrote of emotion as if everyone knows exactly what is meant, but that is not the case. BET has not clarified the concept of emotion nor rendered it a measurable event.

We can point to aspects or components of a typical emotional episode, such as facial expressions, changes in the autonomic nervous system, appraisal of the current situation, overt instrumental behaviour and subjective conscious experience. Psychologists initially attempted to identify emotion with one of these components (James 1884 pointed to conscious experience; Watson 1919 to overt behaviour; and Wenger 1950 to autonomic responses), but no one definition achieved consensus. On another approach, emotion is the package of all the components (Fehr and Russell 1984; Scherer 1984). This package definition is consistent with conventional wisdom: in the Western cultural tradition, we tend to 'see' emotions by packaging together the various components. A similar tendency can be seen when ancient astronomers 'saw'

constellations made up of stars that were actually unrelated cosmologically. These packages make the world seem simpler, serving cognitive economy, but they may be merely convenient fictions. That this is possible for emotion is suggested by the weak correlations found among the components. There is also a logical problem. In this package definition, BET's claim can be re-phrased as the package causing its components, as the whole causing its parts, but such an implication is not stated explicitly in BET (Sias and Bar-On, this volume, discuss a similar criticism of Green's view that emotions are perceptible because their facial expressions are parts of them).

Defining 'emotion' as a package also immediately raises questions about the signalling theory. Consider those cases in which fear and a facial expression of fear co-occur. If fear is a package of components, then the claim that fear causes the fear facial expression raises this question: which component (or combination) is causal? Is it the subjective experience of feeling afraid? Is it changes in the autonomic nervous system? Is it visual scanning of the environment for an escape route? The latter seems a plausible candidate. I anticipate that visual scanning of the environment, whether accompanied by other fear components or not, causes facial changes.

On yet another interpretation, an emotion is an unseen entity behind the components: Fear makes us flee, makes our heart race, makes us feel certain ways and moves our faces. Perhaps, on this interpretation, emotion is Tomkins' affect program, interpreted as a hard-wired brain circuit (Zachar and Ellis 2012). Unfortunately, neuroscientists are abandoning the notion of hard-wired emotion-specific brain circuits (LeDoux 2012, 2014; Lindquist et al. 2012). Of course, the brain is producing all the components of an emotional episode, but the theory that the components of all and only fear episodes, for example, are explained by a specific brain circuit is reminiscent of faculty psychology in which remembering was explained by the memory faculty, imaginings by the imagination faculty and moral behaviour by the morality faculty.

In short, the problems encountered when BET tried to convert everyday concepts of emotion into a scientific theory challenges the scientific usefulness of those concepts. And scrutiny of the concepts themselves raises further questions. The key concepts in BET (anger, fear etc.) originated in folk psychology, concepts that are vaguely defined, heterogeneous, historically shifting, culture-specific and permeated with questionable assumptions. We need to search for alternative conceptual frameworks, and I have been trying to develop one.

In an account I call psychological construction (Russell 2003), I offer an alternative framework for how to think about emotion. People get angry or scared, obviously. But the question is how to develop a scientific account of such episodes. From my viewpoint, the concepts of emotion and of specific emotions must be treated as folk concepts. Their use as scientific concepts has run its course, and it is time to abandon that tack. Thus analyses of our everyday concepts do not reveal the nature of emotion but do reveal a cognitive system used to interpret certain episodes experienced or interpreted as emotional. Such analyses are important, but not for the reasons often offered. Psychological construction abandons the assumption that the folk terms of *emotion*, *anger*, *fear* and the like refer to natural kinds. For that, we must look elsewhere.

Episodes described as emotional consist of changes in various component processes, which are not themselves emotions. No one component is necessary or sufficient to warrant describing the episode as emotional. No clear boundary separates emotional from non-emotional episodes or anger from non-anger episodes. (Instead, all human events involve what I call core affect, although to varying degrees.) No component is caused by the 'emotion'; rather, each component has its own semi-independent process. For example, the autonomic nervous system adjusts in order to enable actions to be performed, whether those actions are part of an emotional episode or not. The components are coordinated, as are all human processes, but not by an emotion brain programme dedicated exclusively to emotion. A witness, scientist or the self might categorize the episode as an emotion, or a specific emotion, but that too is a process to be studied. The component processes are likely universal, although the content might not be (e.g., people might universally categorize their own states, but the cognitive categories used might vary with language and culture). Psychological construction abandons the assumption that emotional episodes are pre-fabricated; it proposes instead that they are assembled in the moment to suit current circumstances.

A psychological constructionist approach implies that facial changes can be accounted for without invoking emotion concepts. (An exception to that generalization might be seen in the invocation of certain display rules; e.g., if a young boy categorizes himself as afraid, but has a display rule that boys never show fear, then he might adjust his demeanour to suggest nonchalance. In this case, 'afraid' is a folk concept that played a role in the boy's thinking.) Some facial changes are

part of perceptual processes: we widen our eyes when scanning the environment (whether surprised or not), and we scrunch our noses when smelling or tasting novel food (whether disgusted or not). Some facial changes are evolutionarily based signals directed at an audience and part of social negotiation: we smile to invite social interaction or collusion (whether happy or not), and we stare to threaten possible retaliation (whether angry or not). Some facial changes are part of paralanguage, and some are part of preparation for behaviour. Society provides rules about appropriate facial behaviour in different situations (downcast eyes during a funeral, whether sad or not), and we try to conform to those rules. Onlookers interpret a person's facial changes, along with vocal and bodily changes, in light of the current situation. Interpretation is based on the folk theory of the mind of others. Interpretation includes (but is not limited to) an emotion categorization. Children initially interpret facial changes in terms of valence (how happy or unhappy the expresser appears to be) and, in Western societies, gradually come to interpret facial changes in terms of discrete emotions (Widen and Russell 2013).

References

- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M. and Bentin, S. 2008. 'Angry, Disgusted, or Afraid? Studies on the Malleability of Emotion Perception', *Psychological Science* 19: 724–732.
- Bateson, G. and Mead, M. 1942. *Balinese Character: A Photographic Analysis*. New York: New York Academy of Sciences.
- Baumeister, R. F., Vohs, K. D., DeWall, C. N. and Zhang, L. 2007. 'How Emotion Shapes Behavior: Feedback, Anticipation, and Reflection, Rather Than Direct Causation', *Personality and Social Psychology Review* 11: 167–203.
- Buss, D. M. 2014. 'Comment: Evolutionary Criteria for Considering an Emotion "Basic": Jealousy as an Illustration', *Emotion Review* 6: 313–315.
- Buzby, D. E. 1924. 'The Interpretation of Facial Expression', *American Journal of Psychology* 35: 602–604.
- Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M. and Ito, T. A. 2000. 'The Psychophysiology of Emotion', *Handbook of Emotions* 2: 173–191.
- Carroll, J. M. and Russell, J. A. 1996. 'Do Facial Expressions Signal Specific Emotions? Judging Emotion from the Face in Context', *Journal of Personality and Social Psychology* 70: 205–218.
- Crivelli, C., Carrera, P. and Fernández-Dols, J. M. 2015. 'Are Smiles a Sign of Happiness? Spontaneous Expressions of Judo Winners', *Evolution and Human Behavior* 36: 52–58.
- Darwin, C. 1872. *The Expression of the Emotions in Man and Animals*. Chicago: University of Chicago Press, 1965.

- DiGirolamo, M. A. and Russell, J. A. 2014. 'The Emotion Seen in a Face as a Methodological Artifact: The Process of Elimination Hypothesis', Ms submitted.
- Ekman, P. 1972. 'Universal and Cultural Differences in Facial Expressions of Emotions', in Cole (ed.) *Nebraska Symposium on Motivation*, 1971. Lincoln: University of Nebraska Press, pp. 207–283.
1980. *The Face of Man: Expressions of Universal Emotions in a New Guinea Village*. New York: Garland Press.
- Ekman, P. and Cordaro D. 2011. 'What Is Meant by Calling Basic Emotions'. *Emotion Review* 3: 364–370.
- Ekman, P. and Friesen, W. V. 1971. 'Constants across Cultures in the Face and Emotion', *Journal of Personality and Social Psychology* 17: 124–129.
- Fehr, B. and Russell, J. A. 1984. 'Concept of Emotion Viewed from a Prototype Perspective', *Journal of Experimental Psychology: General* 113: 464–484.
- Feleky, A. M. 1914. 'The Expression of the Emotions', *Psychological Review* 27: 33–41.
- Fernández-Dols, J. M., Carrera, P. and Crivelli, C. 2011. 'Facial Behavior While Experiencing Sexual Excitement', *Journal of Nonverbal Behavior* 35: 63–71.
- Fernández-Dols, J. M. and Crivelli, C. 2013. 'Emotion and Expression: Naturalistic Studies', *Emotion Review* 5: 24–29.
- Fernández-Dols, J. M. and Ruiz-Belda, M. A. 1995. 'Are Smiles a Sign of Happiness? Gold Medal Winners at the Olympic Games', *Journal of Personality and Social Psychology* 69: 1113–1119.
- Fernández-Dols, J. M., Carrera, P. and Crivelli, C. 2011. 'Facial Behavior While Experiencing Sexual Excitement', *Journal of Nonverbal Behavior* 35: 63–71.
- Fridlund, A. J. 1994. *Human Facial Expression: An Evolutionary View*. San Diego: Academic Press.
- Gendron, M., Roberson, D., van der Vyver, J. M. and Barrett, L. F. 2014. 'Perceptions of Emotion from Facial Expressions Are Not Culturally Universal: Evidence from a Remote Culture', *Emotion* 14: 251–262.
- Green, M. 2007. *Self-Expression*. Oxford: Oxford University Press.
- Griffiths, P. E. 2003. 'Basic Emotions, Complex Emotions, Machiavellian Emotions', *Royal Institute of Philosophy Supplement* 52: 39–67.
- Guilford, J. P. 1929. 'An Experiment in Learning to Read Facial Expression', *Journal of Comparative Psychology* 24: 191–202.
- Izard, C. E. 1971. *The Face of Emotion*. New York: Appleton-Century-Crofts.
- Jack, R. E., Garrod, O. G., Yu, H., Caldara, R. and Schyns, P. G. 2012. 'Facial Expressions of Emotion Are Not Culturally Universal', *Proceedings of the National Academy of Sciences* 109: 7241–7244.
- James, W. 1884. 'What Is an Emotion?' *Mind* 9: 188–205.
- Kayyal, M. H., Widen, S.C. and Russell, J. 2015. 'Context Is More Powerful Than We Think: Contextual Cues Override Facial Cues Even on Valence'. *Emotion* 15(3): 287–291.
- Keltner, D. and Anderson, C. 2000. 'Saving Face for Darwin: The Functions and Uses of Embarrassment', *Current Directions in Psychological Science* 9: 187–192.
- Klineberg, O. 1938. 'Emotional Expression in Chinese Literature', *Journal of Abnormal and Social Psychology* 33: 517–520.

- Kraut, R. E. and Johnston, R. E. 1979. 'Social and Emotional Messages of Smiling: An Ethological Approach', *Journal of Personality and Social Psychology* 37: 1539–1553.
- Krumhuber, E. G. and Manstead, A. S. 2009. 'Can Duchenne Smiles Be Feigned? New Evidence on Felt and False Smiles', *Emotion* 9: 807–820.
- Kunz, M., Prkachin, K. and Lautenbacher, S. 2009. 'The Smile of Pain', *Pain* 145: 273–275.
- Landis, C. 1924. 'Studies of Emotional Reactions: II. General Behavior and Facial Expression', *Journal of Comparative Psychology* 4: 447–509.
- Langfeld, H. S. 1918a. 'The Judgment of Emotions from Facial Expressions', *Journal of Abnormal and Social Psychology* 13: 172–184.
- 1918b. 'Judgments of Facial Expression and Suggestion', *Psychological Review* 25: 488–494.
- LeDoux, J. E. 2012. 'Rethinking the Emotional Brain', *Neuron* 73: 653–676.
2014. 'Afterword: Emotion Construction in the Brain', in Barrett and Russell (eds.) *The Psychological Construction of Emotion*. New York: Guilford, pp. 459–463.
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E. and Barrett, L. F. 2012. 'The Brain Basis of Emotion: A Meta-Analytic Review', *Behavioral and Brain Sciences* 35: 121–143.
- Martens, J. P., Hamlin, J. K. and Tracy, T. L. 2013. 'Is Pride Recognition Innate? Developmental Origins of Understanding the Nonverbal Pride Expression', Poster session presented at the Society for Personality and Social Psychology, New Orleans, LA.
- Nelson, N. L. and Russell, J. A. 2013. 'Universality Revisited', *Emotion Review* 5: 8–15.
2014. 'A Facial Expression of Pax: Assessing Children's "Recognition" of Emotion from Faces'. Manuscript.
- Ortony, A. and Turner, T. J. 1990. 'What's Basic about Basic Emotions?', *Psychological Review* 97: 315–331.
- Parkinson, B. 2005. 'Do Facial Movements Express Emotions or Communicate Motives?', *Personality and Social Psychology Review* 9: 278–311.
- Reisenzein, R. 2000. 'Exploring the Strength of Association between the Components of Emotion Syndromes: The Case of Surprise', *Cognition and Emotion* 14: 1–38.
- Reisenzein, R., Studtmann, M. and Horstmann, G. 2013. 'Coherence between Emotion and Facial Expression: Evidence from Laboratory Experiments', *Emotion Review* 5: 16–23.
- Ruiz-Belda, M. A., Fernández-Dols, J. M., Carrera, P. and Barchard, K. 2003. 'Spontaneous Facial Expressions of Happy Bowlers and Soccer Fans', *Cognition and Emotion* 17: 315–326.
- Russell, J. A. 1991. 'Culture and the Categorizations of Emotion', *Psychological Bulletin* 110: 426–450.
1994. 'Is There Universal Recognition of Emotion from Facial Expressions? A Review of the Cross-Cultural Studies', *Psychological Bulletin* 115: 102–141.
1995. 'Facial Expressions of Emotion: What Lies beyond Minimal Universality?', *Psychological Bulletin* 118: 379–391.

2003. 'Core Affect and the Psychological Construction of Emotion', *Psychological Review* 110: 145–172.
- Scarantino, A. 2015. 'Basic Emotions, Psychological Construction, and the Problem of Variability', in Barrett and Russell (eds.) *The Psychological Construction of Emotion*. New York: Guilford.
- Scherer, K. R. 1984. 'On the Nature and Function of Emotion: A Component Process Approach', in K. R. Scherer and P. Ekman (eds.) *Approaches to Emotion*. Hillsdale, NJ: Erlbaum, pp. 293–318.
- Schneider, K. and Josephs, I. 1991. 'The Expressive and Communicative Functions of Preschool Children's Smiles in an Achievement-Situation', *Journal of Nonverbal Behavior* 15: 185–198.
- Shariff, A. F. and Tracy, J. L. 2011. 'Emotion Expressions: On Signals, Symbols, and Spandrels – A Response to Barrett (2011)', *Psychological Science* 20: 407–408.
- Sherman, M. 1927. 'The Differentiation of Emotional Responses in Infants: II. The Ability of Observers to Judge the Emotional Characteristics of the Crying of Infants and of the Voice of an Adult', *Journal of Comparative Psychology* 7: 335–351.
- Tomkins, S. S. 1962. *Affect, Imagery, Consciousness*, Vol. 1. New York: Springer.
- Tracy, J. L. and Robins, R. W. 2008. 'The Nonverbal Expression of Pride: Evidence for Cross-Cultural Recognition', *Journal of Personality and Social Psychology* 94: 516–530.
- Trauffer, N. M., Widen, S. C. and Russell, J. A. 2013. 'Education and the Attribution of Emotion to Facial Expressions', *Psychological Topics* 22: 237–248.
- Watson, J. B. 1919. *Psychology from the Standpoint of a Behaviorist*. Philadelphia: Lippincott.
- Wenger, M. A. 1950. 'Emotion as Visceral Action: An Extension of Lange's Theory', in Reymert (ed.) *The Second International Symposium on Feelings and Emotions*. New York: McGraw-Hill, pp. 3–10.
- Widen, S. L. and Russell, J. A. 2013. 'Children's Recognition of Disgust in Others', *Psychological Bulletin* 139: 271–299.
- Wierzbicka, A. 1999. *Emotions across Languages and Cultures: Diversity and Universals*. Cambridge: Cambridge University Press.
- Zachar, P. and R. Ellis, R. (eds.) 2012. *Categorical versus Dimensional Models of Affect: A Seminar on the Theories of Panksepp and Russell*. Amsterdam: John Benjamin's Publishing.

8. Cultural Specificities in the Transmission and Decoding of Facial Expressions of Emotion

Rachael E. Jack

As a highly social species, the mutual understanding of emotions is fundamental to the functioning of all human societies. Specifically, communicating emotions – internal states that can influence cognition and behaviour – can provide reliable predictions about the environment, thereby allowing others to adjust their cognitions and behaviours for adaptive action. For example, accurately detecting anger and responding with a submissive behaviour (or a greater display of physical threat) typically avoids a harmful or potentially fatal interaction (Dukas 1998). In contrast, where communication breaks down – for example, due to difficulties expressing or understanding a facial expression – significant negative consequences can arise including social isolation or physical or mental harm (e.g., Baron-Cohen, Leslie and Frith 1985; Hawkley and Cacioppo 2007; LeMoult et al. 2009). Thus, understanding how accurate emotion communication is achieved and why it breaks down is of central importance in facilitating the smooth functioning of all human societies.

As a highly sophisticated social species (e.g., Wilson 2012), humans can communicate a wide range of nuanced emotions, achieved primarily using the face (but see also voice and body posture, e.g., Belin et al. 2008; de Gelder 2009; Sauter et al. 2010) – a powerful communication tool that is equipped with numerous independent muscles that can be combined at different intensities over time (e.g., Drake, Vogl and Mitchell 2010). Consequently, the human face is capable of generating a broad spectrum of complex dynamic patterns, facial expressions, to support emotion communication. Understanding facial expressions of emotion – the origins and evolution, meaning and function – has long been a source of fascination and debate in

fields such as psychology, anthropology and philosophy. One of the most infamous debates has focused on the universality of facial expressions of emotion – that is, on whether facial expressions of the six basic emotions (happiness, surprise, fear, disgust, anger and sadness) are expressed and recognized in the same way across cultures. On the basis of Darwin's ground breaking theory of their biological and evolutionary origins (Darwin [1872] 1999) followed by Ekman's pioneering cross-cultural recognition studies (e.g., Ekman, Sorenson and Friesen 1969), facial expressions of emotion are widely accepted as universal, with six standardized facial expressions considered the 'gold standard' in research across broad fields (e.g., Adolphs et al. 2008; Berenbaum and Oltmanns 1992; Dailey et al. 2010; Furl et al. 2007; Izard 1994; Keltner, Moffitt and Stouthamer-Loeber 1995; Matsumoto and Willingham 2006, 2009; Matsumoto et al. 1986; Morris et al. 1998; Phelps, Ling and Carrasco 2006; Rosenberg, Ekman and Blumenthal 1998; Schyns, Petro and Smith 2007, 2009; Smith and Schyns 2009; Smith et al. 2005; van Rijsbergen and Schyns 2009). Yet, some cultures systematically misinterpret certain facial expressions of basic emotions (e.g., Biehl et al. 1997; Bouher and Carlson 1980; Ducci, Arcuri and Sineshaw 1982; Ekman 1972; Sorenson and Friesen 1969; Ekman et al. 1987, 1969; Elfenbein and Ambady 2003; Elfenbein et al. 2004; Huang et al. 2001; Jack et al. 2009; Kirouac and Doré 1983, 1985; Matsumoto and Ekman 1989; McAndrew 1986; Shioiri et al. 1999) questioning the universality hypothesis and the validity of these specific facial expressions to communicate emotions in different cultures (see Elfenbein and Ambady 2002; Jack 2013; Mesquita and Frijda 1992; Nelson and Russell 2013; Russell 1994 for reviews and meta-analysis).

In an era of rapid globalization and cultural integration facilitated by a rising digital economy, a precise understanding of facial expressions across cultures is increasingly important to support the new social interaction requirements of modern society. This chapter aims to highlight recent research showing cultural specificities in emotion communication, and demonstrate novel methods designed to address both the limitations of traditional research approaches and the complexity of the empirical challenge. To appreciate the relevance of this work, it is useful to first provide a brief historical overview of research on facial expressions of emotion, the main areas of research requiring development and a description of what emotion communication involves.

8.1 A Brief History of Research on Facial Expressions of Emotion

The history of research on facial expressions of emotion is certainly long and varied, stretching at least as far back as Aristotle and spanning several disciplines including philosophy, psychology, medicine, physiology, law and more recently computer science. Here, I will briefly present the highlights most relevant to this chapter (but see Russell 1994 for an excellent summary of the historical development of the field).

Although many other scholars could be credited with the inception of research on facial expression of emotion, Darwin's work is most often cited as the foundation on which modern research and knowledge is based. In his seminal work *The Expression of the Emotions in Man and Animals* (Darwin [1872] 1999), Darwin argued against the commonly held view that God had provided man with complex facial muscles solely for the purposes of communication, by describing the physiological benefits of producing facial expressions of emotion (see also Susskind et al. 2008). For example, a fear facial expression characterized by wide-opened eyes, flared nostrils and open mouth would increase the input of both visual information and oxygen, thereby facilitating the flee response (e.g., by identifying escape routes and increasing muscle function). Similarly, a disgust facial expression characterized by wrinkled nose, lowered brow and tightly squeezed eyes would protect the vulnerable area of the face against the entry of pathogens and therefore disease. Thus, Darwin argued that facial expressions of emotion originally served a physiologically adaptive function that benefitted the expresser before evolving for the purposes of social communication. For example, given their intrinsic signalling (e.g., the high contrast of the eye whites and nose wrinkles are highly detectable by the visual brain) and predictive qualities (e.g., nose wrinkling regularly paired with noxious contaminants), such facial expressions also benefit receivers by providing reliable information about the environment. Since both the production and perception of such facial expressions would increase the chances of survival (e.g., by escaping or avoiding a threat), any traits that enhanced these abilities would be passed on to future generations by natural selection. On the basis of their evolutionary origins, Darwin argued that the facial expressions of emotion used by modern man for social communication are essentially innate behaviours that have retained their original physiologically adaptive movements, and would therefore support the universal recognition of emotions (see also Andrew 1963). To test his theory, Darwin conducted

a cross-cultural recognition study using a set of photographs of selected facial expressions of emotion and reported that all such facial expressions were recognized across several distinct cultures. On the basis of the success and intuitive appeal of Darwin's theory of natural selection and the origins of facial expressions of emotion, notions of a hard-wired universal language of emotion quickly became a working assumption, with the vast majority of subsequent facial expression research disregarding culture as a factor of variance.

Debate about the universality of facial expressions of emotion, or more specifically their socially learned or innate basis, emerged only decades later due to mounting anthropological evidence gathered during the rise of world exploration. Observations and detailed descriptions of stark cultural differences in behaviours once thought to be innate, such as gestures, greeting customs and expressions of emotion (e.g., Briggs 1970; Gorer 1935; Hearn 1894; Holt 1931; Labarre 1947), formed the foundations of a new era of cultural relativism that argued that facial expressions of emotion (amongst others) are culturally learned rather than innate behaviours (e.g., Bruner and Tagiuri 1954). In his seminal work *Pan-Cultural Elements in Facial Displays of Emotion* (Ekman, Sorenson and Friesen 1969), Ekman aimed to address this debate using methods mirroring that of Darwin – a series of cross-cultural recognition studies using hand-selected photographs of six 'pure displays of a single affect' – 'happiness', 'surprise', 'fear', 'disgust-contempt', 'anger' and 'sadness'. Across five distinct literate and pre-literate cultures, Ekman and colleagues showed that all such facial expressions elicited high agreement across cultures, thereby supporting evidence of 'a pan-cultural element in facial affect display' (Ekman, Sorenson and Friesen 1969: 88). Consequently, Ekman and colleagues developed a standardized set of facial expression stimuli that represented six universal basic emotions, 'happiness', 'surprise', 'fear', 'disgust', 'anger' and 'sadness', each described according to the Facial Action Coding System (FACS, Ekman and Friesen 1976) – an objective system that decomposes any facial behaviour into its constituent Action Units (AUs, Ekman and Friesen 1978). For example, 'sadness' is characterized as the AU pattern of Inner Brow Raiser (AU1), Brow Lowerer (AU4) and Lip Corner Depressor (AU15), whereas the 'happiness' AU pattern comprises Cheek Raiser (AU6) and Lip Corner Puller (AU12). See Ekman, Sorenson and Friesen (2002) for prototypes and major variants. As the first tightly controlled and precisely described set of facial expression stimuli, these FACS-coded facial expressions of

emotion quickly became the ‘gold standard’ in research (Adolphs et al. 2008; Berenbaum and Oltmanns 1992; Dailey et al. 2010; Furl et al. 2007; Izard 1994; Keltner, Moffitt and Stouthamer-Loeber 1995; Matsumoto et al. 1986; Matsumoto and Willingham 2006, 2009; Morris et al. 1998; Phelps, Ling and Carrasco 2006; Rosenberg, Ekman and Blumenthal 1998; Schyns, Petro and Smith 2007, 2009; Smith et al. 2005; Smith and Schyns 2009; van Rijsbergen and Schyns 2009), with numerous cross-cultural recognition studies reporting that each elicited above-chance recognition accuracy across cultures in support of the universality hypothesis (e.g., Biehl et al. 1997; Boucher and Carlson 1980; Ducci, Arcuri and Sineshaw 1982; Ekman 1972, 1987; Ekman, Sorenson and Friesen 1969; Elfenbein and Ambady 2003; Elfenbein et al. 2004; Huang et al. 2001; Jack et al. 2009; Kirouac and Doré 1983, 1985; Matsumoto and Ekman 1989; McAndrew 1986; Shioiri et al. 1999). Thus, with the intuitive flavour of Darwin’s work coupled with the first rigorous analysis and representation of facial expressions of emotion, Ekman’s pioneering cross-cultural recognition studies inspired a new generation of empirical work that largely supported the universality hypothesis, thereby marking a new but familiar chapter in the history of facial expression research.

On the basis of the landslide of empirical work showing cross-cultural recognition of these specific AU patterns, facial expressions of emotion were widely considered the universal language of emotion (and still are today by many; see Hwang and Matsumoto, this volume). Yet, a closer look at the reported data and methods used reveal clear cultural differences in the recognition of facial expressions of emotion, thereby questioning the universality hypothesis. I will briefly review some of the main points here, although detailed critiques are well documented (see Elfenbein and Ambady 2002; Jack 2013; Jack and Schyns 2015; Mesquita and Frijda 1992; Nelson and Russell 2013; Russell 1994; Russell, this volume). To appreciate the relevance of these points, it is useful to first step back from the debate and consider what emotion communication using facial expressions involves.

8.2 Emotion Communication Is a System of Information Transmission and Decoding

Like all other forms of communication, such as Morse code, animal mating calls, plant pollination and cellular and bacterial signalling, emotion communication involves the transmissions and decoding of

information. Briefly stated, a sender (here, a human) encodes a message (e.g., 'I feel sad') as a specific form of information (e.g., a facial expression pattern) and transmits it to a receiver (here, another human). On receiving the transmitted information, the receiver decodes the information pattern (here, a facial expression) using their prior knowledge of that pattern to extract the meaning encoded within (e.g., 'he feels sad'). For example, if the downturned mouth and eyes corresponds with the receiver's prior knowledge of sadness (acquired innately or through learning), the receiver will perceive that the sender is sad. Thus, communication is the act of an individual sending information that reliably affects another individual, thereby reflecting a reduction in uncertainty (i.e., the receiver's behaviour reliably predicts the sender's behaviour, e.g., Dukas 1998; Scott-Phillips 2008; Shannon 2001; Slater and Halliday 1994). For example, by transmitting a sad facial expression, the sender elicits in the receiver the perceptual categorization of 'he feels sad'. Understanding any system of communication therefore requires identifying which specific transmitted information patterns elicit a given response in the receiver (see also Jack and Schyns 2015).

At this stage, it is useful to highlight some of the implications and knowledge that can be acquired from analyzing a communication system. First, establishing, maintaining and developing any communication system relies on a number of fundamental conditions. For example, the sender must be capable of transmitting information that the receiver can detect and accurately decode. Where the sender transmits imperceptible information (e.g., detailed nose wrinkling is visible to humans only at short distances; see Smith and Schyns 2009), or unfamiliar information (e.g., Chinese facial expressions of confusion are unfamiliar to Westerners; see Chen et al. 2015) or where sender and receiver do not share the same meaning of the transmitted information (e.g., averted gaze is associated with different emotions across cultures; see Jack, Caldara and Schyns 2011), communication tends to break down. Thus, knowledge of the receiver's detection capabilities (here, the human visual system) and their stored information patterns and associated meanings (i.e., prior knowledge or mental representations of facial expressions) is central to understanding how accurate communication is achieved and why it could break down (e.g., due to a mismatch between transmitted information and the receiver's mental representations). Second, specific definitions of the system of communication also discuss various other requirements (e.g., Dukas 1998; Scott-Phillips 2008), such as transmitted information being intentional (e.g.,

skin melanin pigmentation is transmitted involuntary), specifically evolved for communication (e.g., skin melanin pigmentation evolved as sun protection), conferring an advantage to sender or receiver (e.g., in pigmentocracies, high skin melanin pigmentation disadvantages the sender; see Telles 2014) and sharing honest information (e.g., skin melanin pigmentation is rarely convincingly faked). Such distinctions reflect the intricacies of understanding both the communicatory function (i.e., veridical, involuntary) and phylogenetic and ontogenetic evolutionary bases (i.e., within a culture) of transmitted information patterns. For example, identifying which information patterns are used intentionally, and those that are involuntary, fixed and possibly non-concealable, informs *how* the face is used to optimize success within an ecological niche. In all cases – voluntary or involuntary, advantageous or detrimental, evolved for communication, or dishonest – all such forms of communication rely on the transmission of information that consistently elicits a response in the receiver. Thus, in the case of emotion communication, identifying precisely which facial expression patterns are consistently perceived as specific emotions remains central to addressing the universality debate.

Having described the main components of communication, the transmission and decoding of information and the symbiotic relationship between these components – transmitted information is specifically designed to be detected and accurately decoded by the sender – let us return to our analysis of the cross-cultural recognition studies that are typically cited in support of the universality hypothesis.

8.3 Re-Visiting the Universal Recognition of Facial Expressions of Emotion

Following Ekman's seminal work and the development of a standardized set of so-called universal facial expressions, numerous cross-cultural recognition studies reported that six basic emotions – 'happiness', 'surprise', 'fear', 'disgust', 'anger' and 'sadness' – each represented as a specific AU pattern were recognized across all cultures. Yet, a closer inspection of the reported results reveals systematic and significant cultural differences in recognition performance. Whereas all six facial expressions are recognized with high-recognition accuracy in Western cultures, facial expressions such as 'fear' and 'anger' elicit significantly lower recognition accuracy outside of Western culture, with systematic confusions between emotion categories. For example,

in East Asian cultures ‘fear’ and ‘surprise’ are typically confused, as are ‘disgust’ and ‘anger’. Although in a robust finding, such miscategorizations were largely disregarded in favour of the universality hypothesis primarily due to the use of above-chance performance (>16.7 per cent in a standard six-alternative forced choice task) to demonstrate universal recognition. However, such low levels of accuracy neither reflect typically developed or functional recognition (see Duchaine and Nakayama 2006), nor are sufficiently sensitive to identify significant cultural differences, as demonstrated. Rather, in line with the system of communication as described, low recognition performance – that is, a breakdown in communication – could reflect a mismatch between the transmitted information, the universal facial expression and the cultural receiver’s prior knowledge (i.e., mental representation) of facial expressions of emotion. Although offering some possible explanations, such as language barriers, an empirical investigation and account of these cultural differences remained largely absent, as reflected by Ekman in his book *Emotions Revealed*: ‘To this day I do not know why fear and surprise were not distinguished from each other’ (Ekman 2007: 6). I will now present a series of complementary empirical works that aims to identify the source of the reported cross-cultural miscommunication of facial expressions of emotion by examining the main components of communication – the transmission and decoding of facial expression information.

8.4 Culture-Specific Decoding of Universal Facial Expressions Produces Confusions

In a 2009 research paper in *Current Biology*, Jack and colleagues (2009) examined the decoding strategies used by Westerners and East Asian receivers to categorize the six standardized universal facial expressions of emotion: ‘happiness’, ‘surprise’, ‘fear’, ‘disgust’, ‘anger’ and ‘sadness’. Using a combination of behavioural, spatio-temporal eye movement analysis and statistical modelling revealed that East Asian receivers use a decoding strategy that samples ambiguous eye information from these facial expressions of emotion, giving rise to significant confusions between ‘surprise’ and ‘fear’, and ‘disgust’ and ‘anger’. Specifically, in line with previous cross-cultural recognition studies, Westerners recognized all six facial expressions of emotion with high accuracy, whereas East Asian receivers showed significantly lower recognition performance for ‘fear’ and ‘disgust’, systematically confusing

them with 'surprise' and 'anger', respectively. Using eye-tracking technology to record the fixation patterns of Western and East Asian receivers, a spatio-temporal analysis showed that Westerners distributed their fixations evenly across the face including the eyes and mouth, whereas East Asian receivers tended to concentrate their fixations in the upper part of the face including the eyes. [Figure 8.1](#) shows the results.

Fixation patterns: Each image shows the fixation pattern used by Western and East Asian receivers while categorizing the six basic facial expressions of emotion. Here, all fixation patterns are pooled across six emotions for each culture separately, and displayed on a race-congruent facial expression of fear, for illustrative purposes. Note in the Western group the sampling of information from across the face, including the eyes and mouth, whereas in the East Asian group, the mouth is relatively neglected. **Behavioural confusions:** Here, each matrix shows the pattern of correct and incorrect emotion category responses for each culture, where lighter squares across the diagonal show high accuracy (see colour bar on right) and off-diagonal lighter squares show specific confusions. For example, the East Asian group systematically categorizes 'fear' stimuli as 'surprise', and 'disgust' as 'anger' (see white dashed squares).

Further analysis of the temporal order of fixations showed that East Asian receivers repetitively fixated the left and right eyes of facial expressions that are typically confused – 'surprise', 'fear', 'anger' and 'disgust'. Thus, while Westerns sampled information from the eyes and mouth, East Asian receivers sampled information largely from the eye region. On the basis of the physical similarity of the eye region between 'surprise' and 'fear', and 'disgust' and 'anger', East Asian behavioural confusions could be due to their decoding strategy of fixating the eyes. To test this hypothesis objectively, a model observer sampled information from the same facial expressions and categorized each according to the six emotions. Results showed that by sampling the upper part of the face such as the eyes, the model observer produced a pattern of confusions most similar to that of the East Asian receivers. Thus, the authors concluded that the systematic pattern of East Asian confusions between 'surprise' and 'fear' and 'disgust' and 'anger' is due to a culture-specific decoding strategy that samples ambiguous eye information.

By identifying a source of the East Asian behavioural confusions, several new questions emerged. For example, in line with previous results, standardized universal facial expressions such as 'fear' and 'disgust' are not well recognized in East Asian culture, thereby questioning

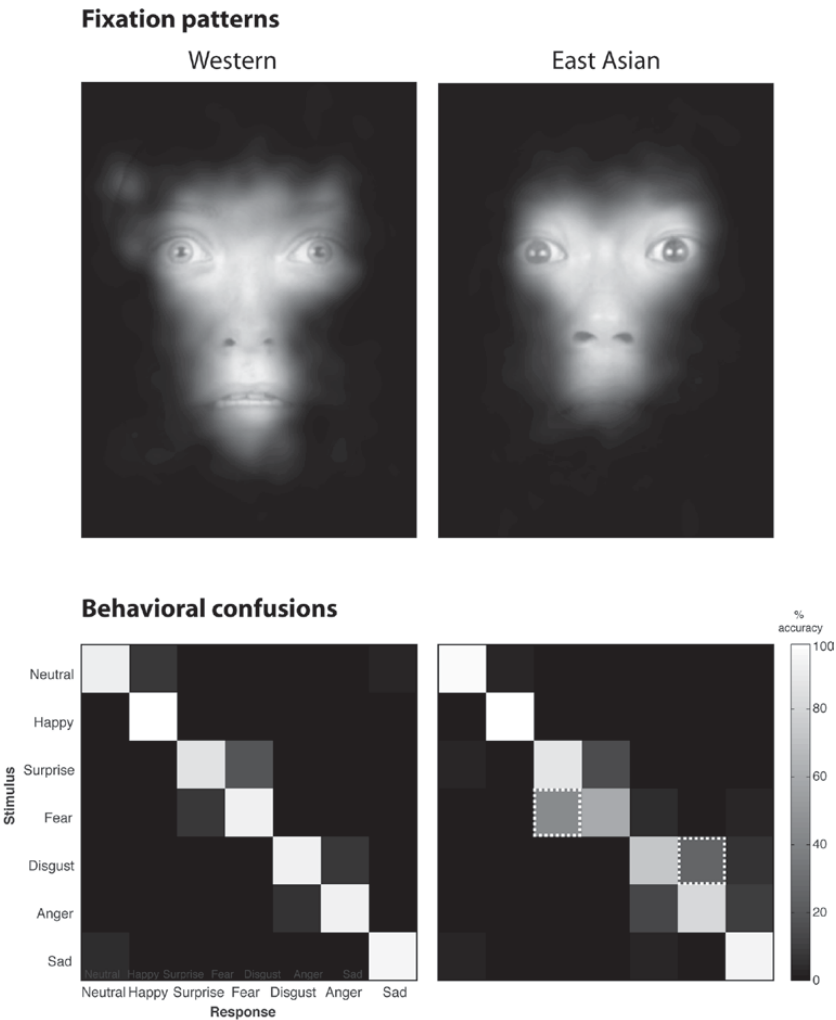


Figure 8.1. Fixation patterns and behavioural confusions when categorizing facial expressions of emotion in different cultures.

Fixation patterns. Each image shows the fixation patterns made by Western and East Asian receivers during the categorization of the six basic emotions. All fixations are pooled across the six emotions and displayed on race-congruent facial expressions of fear, for illustrative purposes. Note that Westerners sampled information from across the face, including the eyes and mouth, whereas East Asian receivers tended to neglect the mouth region while repeatedly sampling the eyes.

Behavioural confusions. Each matrix shows the pattern of correct and incorrect emotion category responses for each culture. Lighter squares across the diagonal show high accuracy (see colour bar to the right), with off-diagonal lighter squares showing specific confusions between emotion categories. For example, East Asian receivers systematically categorize ‘fear’ stimuli as ‘surprise’, and ‘disgust’ as ‘anger’ (see white dashed squares).

their validity to communicate emotion. Therefore, which facial expression patterns do support emotion communication in East Asian culture? Furthermore, eye movements reflect the information sampling strategies used by the receiver to perform a task – that is, the extraction of diagnostic information for accurate recognition. Thus, repetitive sampling of the eye region by East Asian receivers suggests that diagnostic information is expected to be located primarily in the eye region of facial expressions of emotion. To address these questions, Jack and colleagues used a psychophysical approach, reverse correlation (Ahumada and Lovell 1971), to model the perceptual expectations (i.e., mental representations) of facial expressions of the six basic emotions in Western and East Asian receivers (see Dotsch et al. 2008; Gosselin and Schyns 2003; Kontsevich and Tyler 2004 for similar applications of reverse correlation). To illustrate and appreciate the strengths of using reverse correlation techniques, it is useful to return first briefly to the system of communication described.

8.5 Modelling Facial Expressions of Emotion Using a Data-Driven Psychophysics Approach

Understanding any system of communication requires identifying which specific information patterns elicit a given response in the receiver, such as the perceptual category of ‘fear’ in an East Asian receiver. Traditional approaches to addressing this question typically used observational and theoretical knowledge to select a specific set of facial expressions – specific patterns of AUs – and submit them to recognition across cultures. Such an approach, although undoubtedly influential and instrumental in advancing knowledge, has provided a relatively narrow account of how facial expressions communicate emotions within and across cultures for a number of reasons. For example, the selection and testing of only a small set of static facial expressions of emotion coupled with reports and widespread acceptance of universality provided little reason to either explore whether other facial expression patterns communicated basic emotions, or search for those with higher cultural validity (see Jack and Schyns 2015 for a fuller discussion).

In contrast, psychophysical methods use a data-driven approach where the selection of facial expression patterns is largely agnostic (i.e., randomly generated) and the cultural receiver, such as an East Asian, is asked to identify those that correspond to a given emotion, such as

‘fear’. Such an approach has a number of advantages in that no *a priori* assumptions are made about which facial expression patterns will elicit a given response in the receiver, and are therefore sufficiently flexible to capture information patterns that may not be intuitive to the experimenter. Furthermore, recall that on receiving the transmitted information, the receiver uses their prior knowledge (i.e., mental representations) of that pattern to extract the meaning encoded within. Thus, such methods can provide a robust estimate, or model, of the stored information patterns and their associated meanings in individual receivers. Since mental representations are typically developed based on experiences interacting with the environment, thereby providing predictions to guide cognition and behaviour for adaptive action within an ecological niche, capturing such information can inform the information patterns that are transmitted in the receiver’s environment and the receiver’s expectations and likely response. For example, the ‘thumbs up’ gesture means ‘OK’ in Western cultures, but is phallic insult in the Middle East (see Labarre 1947; Morris 1979; Turnbull 1813 for similar examples). To identify which facial expression patterns are associated with the six basic emotions in Western and East Asian culture, Jack and colleagues used the psychophysical technique of reverse correlation coupled with subjective perception (Jack, Caldara and Schyns 2011). Figure 8.2 illustrates the technique.

Stimulus generation and task: On each experimental trial, a template of white noise (i.e., random values per pixel) is added to a neutral base face, thereby creating a perceptively different stimulus. Thus, by adding random values to the base face, the appearance of its features is changed. The receiver, who is naïve to the nature of the stimulus, categorizes the stimulus according to the six basic emotions (e.g., on trial 1, ‘surprise’, on trial n ‘anger’) if the pattern of information correlates with their mental representation (i.e., prior knowledge) of that facial expression, otherwise selecting ‘other’. Thus, where a stimulus is categorized as expressive, the white noise template contains an information pattern that when added to the neutral base face elicits the perception of a specific emotion in the receiver (e.g., ‘surprise’). To obtain a robust estimate of the information patterns associated with each of the six emotion categories, each receiver categorizes a large number of such stimuli.

Analysis: Following the experiment, each of the six emotion categories is associated with a set of white noise templates, each of which contains information patterns that elicit the perception of different emotions. To visualize these information patterns for each emotion and

Stimulus Generation & Task

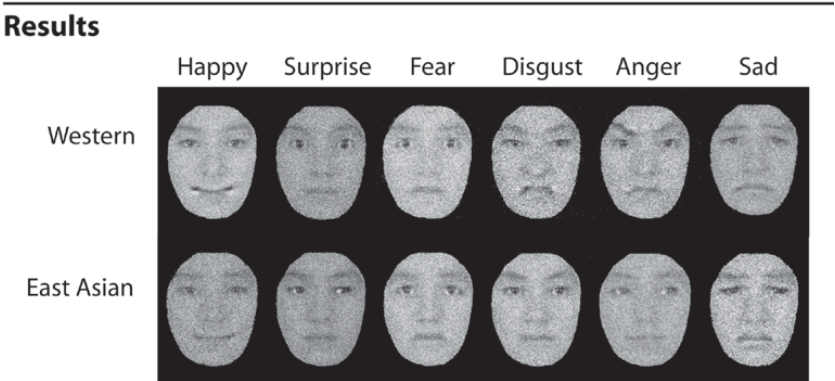
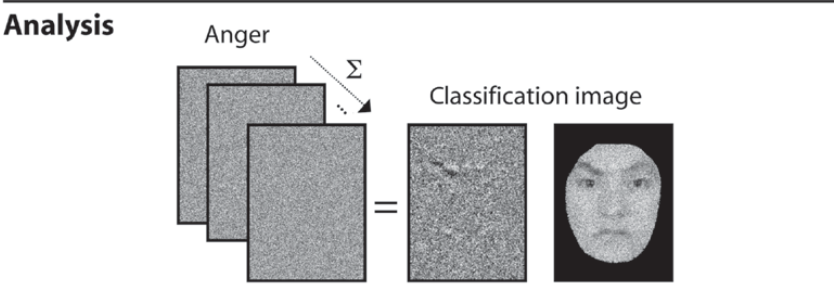
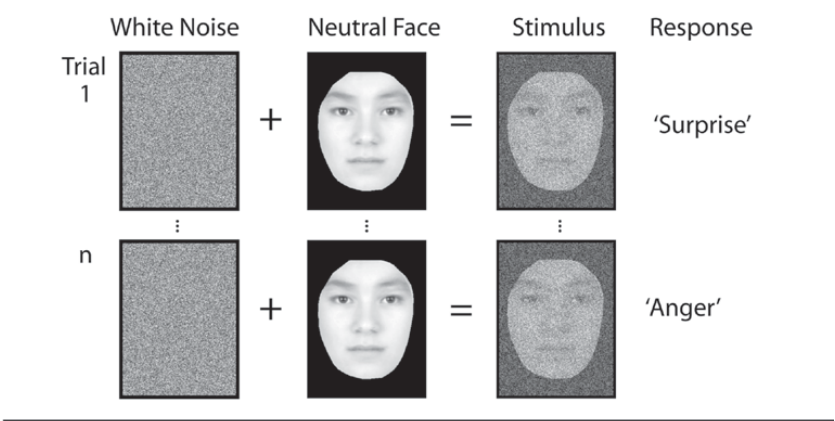


Figure 8.2. Using reverse correlation to model facial expressions of emotions across cultures.

Stimulus generation and task. On each experimental trial, a white noise template is added to a neutral base face to create a perceptively different stimulus. The receiver categorizes the stimulus as expressive (e.g., ‘surprise’) if the information pattern corresponds with their mental representation (i.e., perceptual expectations) of that facial expression. To obtain a robust estimate of the information patterns required for the perception of each emotion, each receiver completes a high number of such trials. *Analysis.* Each emotion is associated with a set of white noise templates, each

Figure 8.2 (*cont.*)

of which contains patterns of information that when added to the neutral base face elicits the perception of a given emotion in the individual receiver. To visualize these information patterns, the white noise templates are summed to create a classification image and then added to the neutral base face. All such resulting facial expressions are then validated by another group of receivers in each culture. *Results.* Individual receiver results from Western and East Asian culture are shown here. Statistical analysis of the classification images shows that Westerners needed information added to the eyes and mouth to perceive different emotions, whereas East Asian receivers needed information added primarily to the eyes. Note, for example, in the East Asian group the distinct changes in gaze direction and relative invariance of the mouth.

receiver separately, each set of white noise templates is summed to produce a classification image and then added to the neutral base face. For example, to elicit the perception of 'anger' in this individual receiver, a lowered brow and downturned mouth are required, as can be seen from both the classification image and its application to the neutral base face. Using this technique, a group of Western and East Asian receivers each completed the task, thereby producing a set of culture-specific mental representations of the six basic facial expressions of emotion computed for each individual receiver. Each resulting facial expression was then validated using a separate group of receivers from the same culture.

Results: To illustrate the main cultural differences, individual receiver results are shown here (see Jack, Caldara and Schyns 2011 for full details of analysis and results). Statistical analysis of the resulting classification images showed that Westerners required information added to the eyes and mouth to perceive different emotions. For example, note how the eyes, eyebrows and mouth each varies across the six emotion categories. In contrast, East Asian receivers required information to be added primarily to the eye region. Note, across the emotion categories the distinct changes in gaze direction and relative invariance of the mouth. Interestingly, our results mirror the information patterns used to communicate emotions in cyberspace – emoticons. Western emoticons, such as :) for 'happy' and :(for 'sad' tend to vary around the mouth, whereas East Asian emoticons, such as (^.^) for 'happy' and (>.<) for 'angry' tend to vary around the eyes to communicate emotions (see also Yuki et al. 2007 for similar cultural differences).

Here, Jack and colleagues showed that receivers from Western and East Asian culture have different expectations about the location and form of diagnostic information for recognizing facial expressions of emotion. These data reveal a further and related source of confusion

during cross-cultural communication – the expectation of which information communicates emotion. For example, East Asian receivers expect diagnostic expressive information to be located in the eye region, reflecting that emotion is communicated primarily with the eye region in East Asian culture. Such expectations in the location of diagnostic information are mirrored by the fixation patterns used when decoding facial expressions of emotion (see fixation patterns in [Figure 8.1](#)). Thus, since prior knowledge can guide an individual's cognition and behaviour such as information sampling through eye movements, cultural differences in expectations can produce confusions at the level of information sampling and the interpretation of that information.

Together, both the preceding eye tracking and reverse correlation data show two different but closely related sources of confusion in the communication of emotions across cultures – expectations about which specific information patterns communicate emotions, and how that information is sampled and decoded by the receiver. Thus, as highlighted earlier, exploration of the receiver's decoding capabilities – their information sampling strategies, stored information patterns and associated meanings – can inform how accurate communication is achieved and why it could break down (e.g., due to a mismatch between transmitted information and the receiver's mental representations).

8.6 Modelling Dynamic Facial Expressions of Emotion across Cultures

A further limitation of the traditional approaches used to understand facial expressions of emotion is the omission of dynamic information. As is clear from any social interaction, facial expressions are characteristically dynamic, some involving rapid, fleeting or sustained activations of face movements that unfold over time as a specific sequence (e.g., Jack, Garrod and Schyns 2014). Indeed, dynamic information such as the acceleration of a specific face movement, or the duration of its activation, plays a central role in the perception of various socially relevant factors, including honesty (e.g., genuine vs. fake smiles, Krumhuber and Kappas 2005) and trustworthiness of the transmitter (Krumhuber et al. 2007; see Krumhuber, Kappas and Manstead 2013 for a review). Thus, providing a precise and authentic characterization of facial expressions of emotion within and across cultures involves extending the exploration of facial expression information to their dynamic features. Using the same reverse correlation approach described combined

with a dynamic facial expression generator (Yu, Garrod and Schyns 2012), Jack and colleagues identified the specific dynamic facial expression patterns that communicate emotions in Western and East Asian culture, revealing cultural differences. Figure 8.3 illustrates the stimulus generation and task procedure.

Stimulus generation and task: On each experimental trial, a dynamic facial expression generator – the Generative Face Grammar (GFG, Yu, Garrod and Schyns 2012) – produces a photorealistic random facial animation by combining a randomly selected subset of AUs (here, Upper Lid Raiser – AU5, Nose Wrinkler – AU9 and Upper Lip Raiser – AU10) each of which is ascribed a random movement (see curves beneath each AU). The receiver categorizes the stimulus according to the six basic emotions (here, ‘disgust’) and rates the intensity of the emotion perceived (here, ‘strong’) if the dynamic AU pattern corresponds to the receiver’s mental representation of that facial expression of emotion at that intensity. If the dynamic AU pattern does not correspond to any of the response categories provided, the receiver selects ‘other’. Following the experiment in which each receiver categorizes a large number of such trials, statistical analyses of the relationship between the dynamic AU patterns and receiver responses reveals the specific facial expression pattern that communicates each emotion. Applying this method in different cultures and comparing the resulting dynamic AU patterns can therefore reveal precisely which face movements could give rise to cultural confusions. Figure 8.4 shows the some of the main results.

Diagnostic emotion information: Here, two facial expressions of emotion – ‘anger’ and ‘disgust’ – are used to illustrate a cultural difference in the location of diagnostic expressive information. In line with previous results, in the Western group, the mouth is more informative in distinguishing these facial expressions (note the difference in the individual face images, and corresponding lighter areas in the difference map). In the East Asian group, the eyes are more informative (note the narrow eyes in ‘disgust’ compared to wide-open eye whites in ‘anger’, as shown by the light areas in the difference map). Analysis of the temporal dynamics of facial expressions of emotion also showed cultural differences in how the face communicates emotional intensity. East Asian facial expressions communicate emotional intensity primarily using early eye region activity. Together, these results support previous work showing cultural differences in the transmission and decoding of facial expressions of emotion, whereby Western and East Asian

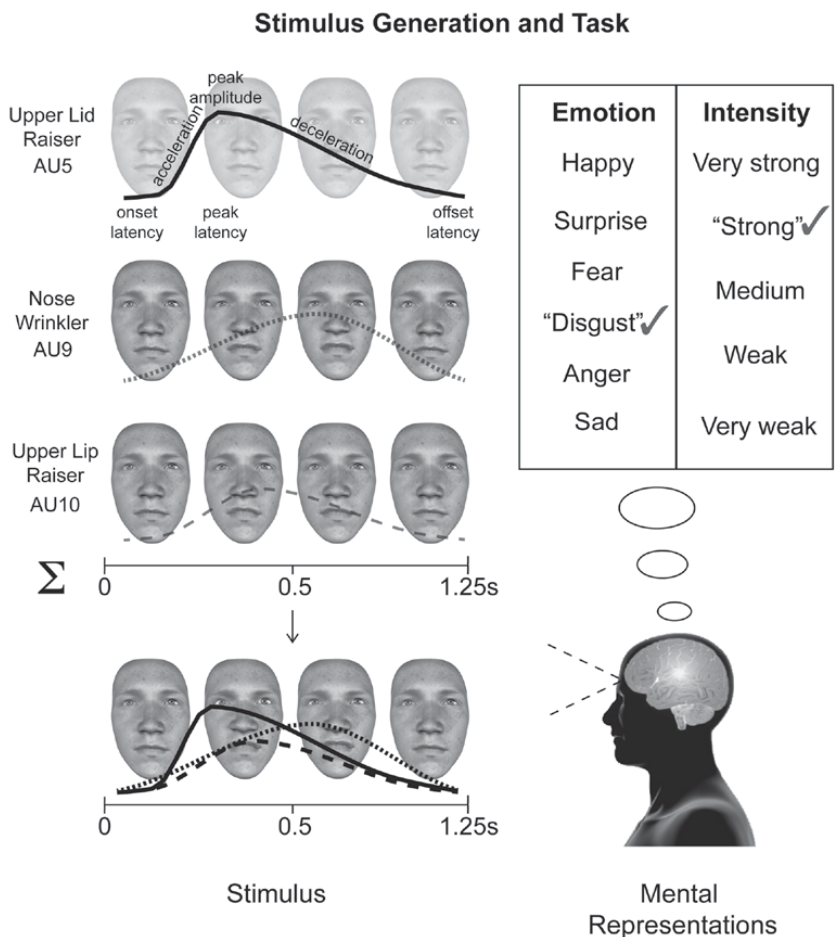


Figure 8.3. Modelling dynamic facial expressions of emotion across cultures. *Stimulus:* On each experimental trial, a computer graphics facial expression generator – the Generative Face Grammar (GFG, Yu, Garrod and Schyns 2012) – randomly selects a subset of AUs from a core set of forty AUs and ascribes a random movement to each (see temporal curves for each AU). The GFG then combines the randomly activated AUs to produce a random facial animation, illustrated here with four snapshots across time. *Mental Representations:* The cultural receiver views the stimulus and categorizes it according to one of the six basic emotions (here, 'disgust') and rates the intensity perceived (here, 'strong') if the dynamic pattern corresponds to their mental representation of that facial expression of emotion and intensity. Otherwise, the receiver selects 'other'. Each receiver completes a large number of such facial animations, presented on different identities.

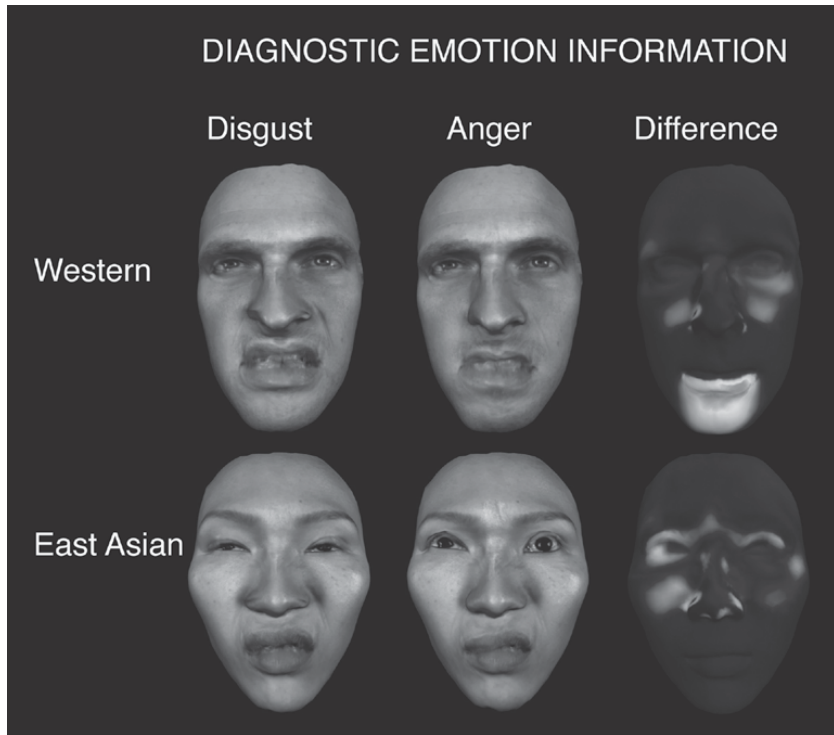


Figure 8.4. Diagnostic facial expression information across cultures.

Two facial expressions of emotion – ‘disgust’ and ‘anger’ – are used here to illustrate a cultural difference in the location of diagnostic emotion information. In the Western facial expressions of ‘disgust’ and ‘anger’, the mouth is more informative to distinguish them. For example, note the difference in the shape of the mouth between the two face images, which is reflected by the lighter areas in the different face image. In contrast, in East Asian facial expressions of ‘disgust’ and ‘anger’, the eyes are more informative. For example, note the tightly squeezed eyes in ‘disgust’ compared to the salient eye whites in ‘anger’, as reflected in the different face image.

culture receivers expect and sample expressive information from different locations of the face.

8.7 Summary

Facial expressions of six basic emotions have largely been considered universal on the basis of their biological and evolutionary origins proposed by Darwin in his early seminal work. Although several decades of subsequent cross-cultural recognition studies reported universal

recognition of a set of standardized facial expressions – specific AU patterns selected based on observational and theoretical knowledge – closer scrutiny of the data reveals marked cultural difference in recognition performance, questioning the universality hypothesis and the validity of these specific facial expression patterns to communicate emotions across cultures. Here, I present converging evidence from a set of complementary studies that aims to address the universality hypothesis debate using novel methods designed to overcome some of the limitations of traditional approaches. Specifically, using the psychophysical technique of reverse correlation, I show that Western and East Asian receivers expect diagnostic expressive information to be located in different regions of the face. Westerners expect diagnostic information to be distributed across the face including the eyes and mouth, whereas East Asian receivers primarily expect the eye region to contain diagnostic information. Such expectations are mirrored by the fixation patterns used by receivers to sample information from the face to categorize facial expressions of emotion. Since the location and form of diagnostic expressive information differs across cultures – for example, East Asian receivers expect changes of gaze direction for emotions such as ‘fear’, ‘anger’ and ‘sad’, whereas Westerners do not – the mismatch in expectations and subsequent information sampling can give rise to cross-cultural confusions. Our results highlight the importance of understanding the cultural receiver’s decoding capabilities, such as their prior knowledge and information sampling strategies, to precisely identify the sources that support accurate communication and those that produce misunderstanding.

I also demonstrate the strengths of using a data-driven psychophysics approach to understanding emotion communication across cultures. Traditional approaches typically selected specific static facial expressions based on theoretical knowledge or observations and used above-chance performance (>16.7 per cent in a typical six-alternative forced choice task) to support claims of universal recognition. However, in addition to above-chance accuracy masking any cultural differences in recognition performance, the facial expression patterns reported as universal elicit significantly different responses across cultures, questioning their validity. Yet, widespread acceptance of these facial expressions as universal and the ‘gold standard’ in research provided little impetus to explore other AU patterns and identify those that accurately communicate basic emotions outside of Western culture. However, such an endeavour represents a genuine empirical challenge, because the face is

capable of producing a broad spectrum of complex dynamic patterns, each of which could potentially communicate an emotion in a specific culture. To illustrate, consider the high combinatorics of all biologically plausible AU patterns, where each AU in that pattern can differ according to several temporal features such as intensity, when in time the AU peaks or its acceleration (e.g., see temporal features illustrated by the curves in [Figure 8.3](#)). Exploring the high dimensional information space of the face – that is, selecting and testing these different dynamic AU patterns in different cultures – requires an objective and systematic approach such as psychophysics. Of course, although the field of psychophysics – one of the oldest disciplines in psychology – far preceded the era of recognition studies pioneered by Ekman, its application to facial expression research at that time would have been far too impractical. For example, generating all dynamic AU combinations using human posers and presented stimuli manually would have created a disproportionate challenge to advancing knowledge at that time. Now, as illustrated earlier, with the development of new technologies such as 3D computer graphics and photorealistic structural face capture systems, coupled with an increase in cross-disciplinary research, modern data-driven approaches can be used to refine both empirical evidence and theoretical developments.

8.8 Future Research Directions

As with many other forms of communication, facial expressions of emotion are typically interpreted within specific contexts, where the same dynamic patterns (e.g., tightly squeezed eyes, bared teeth) could convey different meanings (e.g., winning, losing, pleasure or pain) when combined with additional sources of information (e.g., body movements, vocalizations, or environment. See Barrett, Mesquita and Gendron [2011](#); Hassin, Aviezer and Bentin [2013](#) for reviews; Parkinson [2013](#)). Thus, emotion communication can involve the integration of information from multiple sources to form a hyper-complex information pattern that elicits a specific response in the receiver. Here, we have demonstrated the strengths and benefits of using a psychophysics approach to identify which information patterns communicate specific emotions. However, applying such methods to multiple sources of information presents a genuinely ambitious empirical feat because each source of information – vocalizations, language, body posture, movements or shape, clothing and

hairstyles, environment and so on – represents an explicit hypothesis on the nature of communication. To address this challenge, both empirical and theoretical knowledge can be used to guide the selection and exploration of each information source either separately or in combination to build a rich model of emotion communication in different cultures.

Finally, with the rise of a new digital economy that facilitates globalization and cultural integration, it is increasingly important to equip new interactive technologies, such as social avatars and companion robots, with detailed knowledge of human social communication. By using psychophysical methods to precisely identify the information patterns that humans associate with different social categories, such as emotions as well as personality traits (e.g., trustworthiness, dominance, Dotsch et al. 2008; Gill et al. 2014; Oosterhof and Todorov 2008) and mental states (e.g., thinking, bored, confused, doubtful, Baron-Cohen 1996; Ekman 1979; Fridlund 2014), such knowledge can be installed and used by digital and robotic agents to recognize and adaptively display culture-appropriate social information. For example, companion robots could flexibly adjust their facial movements to suit a specific role, context or culture, such as displaying positive emotions when greeting guests.

Understanding how facial expressions – complex dynamic patterns of information – communicate emotions is a genuine empirical challenge that requires the strategic interdisciplinary convergence of knowledge from broader fields such as vision science and psychophysics, cognitive psychology, cultural and social psychology, anthropology, biology and engineering. Now, with the emergence of a scientific culture that embraces interdisciplinary approaches, the field of facial expression research has reached a new chapter of empirical progress and theoretical developments.

References

- Adolphs, R., Spezio, M. L., Parlier, M. and Piven, J. 2008. 'Distinct Face-Processing Strategies in Parents of Autistic Children', *Current Biology* 18: 1090–1093.
- Ahumada, A. and Lovell, J. 1971. 'Stimulus Features in Signal Detection', *Journal of the Acoustical Society of America* 49: 1751–1756.
- Andrew, R. J. 1963. 'Evolution of Facial Expression', *Science* 142: 1034–1041.
- Baron-Cohen, S. 1996. 'Reading the Mind in the Face: A Cross-Cultural and Developmental Study', *Visual Cognition* 3: 39–60.

- Baron-Cohen, S., Leslie, A. M. and Frith, U. 1985. 'Does the Autistic Child Have a "Theory of Mind"?', *Cognition* 21: 37–46.
- Barrett, L. F., Mesquita, B. and Gendron, M. 2011. 'Context in Emotion Perception', *Current Directions in Psychological Science* 20: 286–290.
- Belin, P., Fillion-Bilodeau, S. and Gosselin, F. 2008. 'The Montreal Affective Voices: A Validated Set of Nonverbal Affect Bursts for Research on Auditory Affective Processing', *Behavior Research Methods* 40: 531–539.
- Berenbaum, H. and Oltmanns, T. F. 1992. 'Emotional Experience and Expression in Schizophrenia and Depression', *Journal of Abnormal Psychology* 101: 37–44.
- Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T. and Ton, V. 1997. 'Matsumoto and Ekman's Japanese and Caucasian Facial Expressions of Emotion (JACFEE): Reliability Data and Cross-National Differences', *Journal of Nonverbal Behavior* 21: 3–21.
- Boucher, J. D. and Carlson, G. E. 1980. 'Recognition of Facial Expression in Three Cultures', *Journal of Cross-Cultural Psychology* 11: 263–280.
- Briggs, J. 1970. *Never in Anger: Portrait of an Eskimo Family*. Cambridge, MA: Harvard University Press.
- Bruner, J. S. and Tagiuri, R. 1954. *The Perception of People*. DTIC Document.
- Chen, C., Garrod, O., Schyns, P. and Jack, R. 2015. 'The Face Is the Mirror of the Cultural Mind', *Journal of Vision* 15(12): 928–928.
- Dailey, M. N., Joyce, C., Lyons, M. J., Kamachi, M., Ishi, H., Gyoba, J. and Cottrell, G. W. 2010. 'Evidence and a Computational Explanation of Cultural Differences in Facial Expression Recognition', *Emotion* 10: 874–893.
- Darwin, C. [1872] 1999. *The Expression of the Emotions in Man and Animals*, 3rd edn. London: Fontana Press.
- de Gelder, B. 2009. 'Why Bodies? Twelve Reasons for Including Bodily Expressions in Affective Neuroscience', *Philosophical Transactions of the Royal Society B: Biological Sciences* 364: 3475–3484.
- Dotsch, R., Wigboldus, D. H., Langner, O. and van Knippenberg, A. 2008. 'Ethnic Out-Group Faces Are Biased in the Prejudiced Mind', *Psychological Science* 19: 978–980.
- Drake, R. L., Vogl, A. and Mitchell, A. W. 2010. *Gray's Anatomy for Students*. Philadelphia: Churchill Livingstone.
- Ducci, L., Arcuri, L. and Sineshaw, T. 1982. 'Emotion Recognition in Ethiopia: The Effect of Familiarity with Western Culture on Accuracy of Recognition', *Journal of Cross-Cultural Psychology* 13: 340–351.
- Duchaine, B. C. and Nakayama, K. 2006. 'Developmental Prosopagnosia: A Window to Content-Specific Face Processing', *Current Opinion in Neurobiology* 16: 166–173.
- Dukas, R. 1998. *Cognitive Ecology: The Evolutionary Ecology of Information Processing and Decision Making*. Chicago: University of Chicago Press.
- Ekman, P. 1972. 'Universal and Cultural Differences in Facial Expression of Emotion', in J. R. Cole (ed.) *Nebraska Symposium on Motivation*, 1971, Vol. 19. Lincoln: Nebraska University Press, pp. 207–283.
1979. 'About Brows: Emotional and Conversational Signals', in Cranach, Foppa, Lepenies and Ploog (eds.) *Human Ethology*. Cambridge: Cambridge University Press, pp. 169–202.

2007. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. New York: Holt Paperbacks.
- Ekman, P. and Friesen, W. V. 1976. 'Measuring Facial Movement', *Journal of Non-verbal Behaviour* 1: 1573–3653.
1978. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., Friesen, W. V. and Hager, J. C. 2002. *Facial Action Coding System Investigator's Guide*. Salt Lake City, UT: Research Nexus.
- Matsumoto, D., Consolacion, T., Yamada, H., Suzuki, R., Franklin, B., Paul, S. et al. 2002. 'American-Japanese Cultural Differences in Judgements of Emotional Expressions of Different Intensities', *Cognition and Emotion* 16: 721–747.
- Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., Ricci-Bitti, P. E., Scherer, K., Tomita, M. and Tzavaras, A. 1987. 'Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion', *Journal of Personality and Social Psychology* 53: 712–717.
- Elfenbein, H. A. and Ambady, N. 2002. 'On the Universality and Cultural Specificity of Emotion Recognition: A Meta-Analysis', *Psychological Bulletin* 128: 203–235.
2003. 'When Familiarity Breeds Accuracy: Cultural Exposure and Facial Emotion Recognition', *Journal of Personality and Social Psychology* 85(2): 276–290.
- Elfenbein, H. A., Mandal, M., Ambday, N., Harizuka, S. and Kumar, S. 2004. 'Hemifacial Differences in the In-Group Advantage in Emotion Recognition', *Cognition and Emotion* 18: 613–629.
- Fridlund, A. J. 2014. *Human Facial Expression: An Evolutionary View*. New York: Academic Press.
- Furl, N., Van Rijsbergen, N. J., Treves, A., Friston, K. J. and Dolan, R. J. 2007. 'Experience-Dependent Coding of Facial Expression in Superior Temporal Sulcus', *Proceedings of the National Academy of Sciences* 104: 13485–13489.
- Gill, D., Garrod, O. G., Jack, R. E. and Schyns, P. G. 2014. 'Facial Movements Strategically Camouflage Involuntary Social Signals of Face Morphology', *Psychological Science* 25: 1079–1086.
- Gorer, G. 1935. *Africa Dances*. New York: W. W. Norton & Co.
- Gosselin, F. and Schyns, P. G. 2003. 'Superstitious Perceptions Reveal Properties of Internal Representations', *Psychological Science* 14: 505–509.
- Hassin, R. R., Aviezer, H. and Bentin, S. 2013. 'Inherently Ambiguous: Facial Expressions of Emotions, in Context', *Emotion Review* 5: 60–65.
- Hawkey, L. C. and Cacioppo, J. T. 2007. 'Aging and Loneliness: Downhill Quickly?' *Current Directions in Psychological Science* 16: 187–191.
- Hearn, L. 1894. *Glimpses of Unfamiliar Japan*. New York: Houghton Mifflin & Co.
- Holt, E. B. 1931. *Animal Drive and the Learning Process*. New York: Henry Holt and Co.
- Huang, Y., Tang, S., Helmeste, D., Shioiri, T. and Someya, T. 2001. 'Differential Judgement of Static Facial Expressions of Emotions in Three Cultures', *Psychiatry and Clinical Neurosciences* 55: 479–483.

- Izard, C. E. 1994. 'Innate and Universal Facial Expressions: Evidence from Developmental and Cross-Cultural Research', *Psychological Bulletin* 115: 288–299.
- Jack, R. E. 2013. 'Culture and Facial Expressions of Emotion', *Visual Cognition* 21: 1–39.
- Jack, R. E. and Schyns, P. G. 2015. 'The Human Face as a Dynamic Tool for Social Communication', *Current Biology* 25: R621–R634.
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G. and Caldara, R. 2009. 'Cultural Confusions Show that Facial Expressions Are Not Universal', *Current Biology* 19: 1543–1548.
- Jack, R. E., Caldara, R. and Schyns, P. G. 2011. 'Internal Representations Reveal Cultural Diversity in Expectations of Facial Expressions of Emotion', *Journal of Experimental Psychology* 141: 19–25.
- Jack, R. E., Garrod, O. G. and Schyns, P. G. 2014. 'Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals over Time', *Current Biology* 24: 187–192.
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G. and Caldara, R. 2009. 'Cultural Confusions Show that Facial Expressions Are Not Universal', *Current Biology* 19: 1543–1548.
- Keltner, D., Moffitt, T. E. and Stouthamer-Loeber, M. 1995. 'Facial Expressions of Emotion and Psychopathology in Adolescent Boys', *Journal of Abnormal Psychology* 104: 644–652.
- Kirouac, G. and Doré, F. Y. 1983. 'Accuracy and Latency of Judgment of Facial Expressions of Emotions', *Perceptual and Motor Skills* 57: 683–686.
1985. 'Accuracy of the Judgment of Facial Expression of Emotions as a Function of Sex and Level of Education', *Journal of Nonverbal Behavior* 9: 3–7.
- Kontsevich, L. L. and Tyler, C. W. 2004 'What Makes Mona Lisa Smile?', *Vision Research* 44: 1493–1498.
- Krumhuber, E. and Kappas, A. 2005. 'Moving Smiles: The Role of Dynamic Components for the Perception of the Genuineness of Smiles', *Journal of Nonverbal Behavior* 29: 3–24.
- Krumhuber, E. G., Kappas, A. and Manstead, A. S. 2013. 'Effects of Dynamic Aspects of Facial Expressions: A Review', *Emotion Review* 5: 41–46.
- Krumhuber, E., Manstead, A. S., Cosker, D., Marshall, D., Rosin, P. L. and Kappas, A. 2007. 'Facial Dynamics as Indicators of Trustworthiness and Cooperative Behavior', *Emotion* 7: 730–735.
- Labarre, W. 1947. 'The Cultural Basis of Emotions and Gestures', *Journal of Personality* 16: 49–68.
- LeMoult, J., Joormann, J., Sherdell, L., Wright, Y. and Gotlib, I. H. 2009. 'Identification of Emotional Facial Expressions Following Recovery from Depression', *Journal of Abnormal Psychology* 118: 828–833.
- Matsumoto, D. and Ekman, P. 1989. 'American-Japanese Cultural Differences in Intensity Ratings of Facial Expressions of Emotion', *Motivation and Emotion* 13: 143–157.
- Matsumoto, D. and Willingham, B. 2006. 'The Thrill of Victory and the Agony of Defeat: Spontaneous Expressions of Medal Winners of the 2004 Athens Olympic Games', *Journal of Personality and Social Psychology* 91: 568–581.

2009. 'Spontaneous Facial Expressions of Emotion of Congenitally and Noncongenitally Blind Individuals', *Journal of Personality and Social Psychology* 96: 1–10.
- Matsumoto, D., Haan, N., Yabrove, G., Theodorou, P. and Carney, C. C. 1986. 'Preschoolers' Moral Actions and Emotions in Prisoner's Dilemma', *Developmental Psychology* 22: 663–670.
- Matsumoto, D. and Willingham, B. 2006. 'The Thrill of Victory and the Agony of Defeat: Spontaneous Expressions of Medal Winners of the 2004 Athens Olympic Games', *Journal of Personality and Social Psychology* 91: 568–581.
- McAndrew, F. T. 1986. 'A Cross-Cultural Study of Recognition Thresholds for Facial Expressions of Emotion', *Journal of Cross-Cultural Psychology* 17: 211–224.
- Mesquita, B. and Frijda, N. H. 1992. 'Cultural Variations in Emotions: A Review', *Psychological Bulletin* 112: 179–204.
- Morris, D. 1979. *Gestures, Their Origins and Distribution*, New York: Stein and Day.
- Morris, J. S., Friston, K. J., Buchel, C., Frith, C. D., Young, A. W., Calder, A. J. and Dolan, R. J. 1998. 'A Neuromodulatory Role for the Human Amygdala in Processing Emotional Facial Expressions', *Brain* 121: 47–57.
- Nelson, N. L. and Russell, J. A. 2013. 'Universality Revisited', *Emotion Review* 5: 8–15.
- Oosterhof, N. N. and Todorov, A. 2008. 'The Functional Basis of Face Evaluation', *Proceedings of the National Academy of Sciences* 105: 11087–11092.
- Parkinson, B. 2013. 'Contextualizing Facial Activity', *Emotion Review* 5: 97–103.
- Phelps, E. A., Ling, S. and Carrasco, M. 2006. 'Emotion Facilitates Perception and Potentiates the Perceptual Benefits of Attention', *Psychological Science* 17: 292–299.
- Rosenberg, E. L., Ekman, P. and Blumenthal, J. A. 1998. 'Facial Expression and the Affective Component of Cynical Hostility in Male Coronary Heart Disease Patients', *Health Psychology* 17: 376–380.
- Russell, J. A. 1994. 'Is There Universal Recognition of Emotion from Facial Expression? A Review of the Cross-Cultural Studies', *Psychological Bulletin* 115: 102–141.
- Sauter, D. A., Eisner, F., Ekman, P. and Scott, S. K. 2010. 'Cross-Cultural Recognition of Basic Emotions through Nonverbal Emotional Vocalizations', *Proceedings of the National Academy of Sciences of the United States of America* 107: 2408–2412.
- Schyns, P. G., Petro, L. S. and Smith, M. L. 2007. 'Dynamics of Visual Information Integration in the Brain for Categorizing Facial Expressions', *Current Biology* 17: 1580–1585.
2009. 'Transmission of Facial Expressions of Emotion Co-Evolved with Their Efficient Decoding in the Brain: Behavioral and Brain Evidence', *PLoS One* 4: e5625.
- Scott-Phillips, T. C. 2008. 'Defining Biological Communication', *Journal of Evolutionary Biology* 21: 387–395.
- Shannon, C. E. 2001. 'A Mathematical Theory of Communication', *ACM SIGMOBILE Mobile Computing and Communications Review* 5: 3–55.

- Shioiri, T., Someya, T., Helmeste, D. and Tang, S. W. 1999 'Misinterpretation of Facial Expression: A Cross-Cultural Study', *Psychiatry and Clinical Neurosciences* 53: 45–50.
- Slater, P. J. and Halliday, T. R. 1994. *Behaviour and Evolution*. Cambridge: Cambridge University Press.
- Smith, F. W. and Schyns, P. G. 2009. 'Smile through Your Fear and Sadness', *Psychological Science* 20: 1202–1208.
- Smith, M. L., Cottrell, G. W., Gosselin, F. and Schyns, P. G. 2005. 'Transmitting and Decoding Facial Expressions', *Psychological Science* 16: 184–189.
- Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W. and Anderson, A. K. 2008. 'Expressing Fear Enhances Sensory Acquisition', *Nature Neuroscience* 11: 843–850.
- Telles, E. 2014. *Pigmentocracies: Ethnicity, Race, and Color in Latin America*. Chapel Hill: University of North Carolina Press Books.
- Turnbull, J. 1813. *A Voyage Round the World in the Years 1800, 1801, 1802, 1803, and 1804 in which the Author Visited the Principle Islands in the Pacific Ocean and the English Settlements of Port Jackson and Norfolk Island*, London: Printed for R. Phillips by T. Gillet, 1805.
- van Rijsbergen, N. J. and Schyns, P. G. 2009. 'Dynamics of Trimming the Content of Face Representations for Categorization in the Brain', *PLoS Computational Biology* 5: e1000561.
- Wilson, E. O. 2012. *The Social Conquest of Earth*. London: W. W. Norton & Co.
- Yu, H., Garrod, O. G. B. and Schyns, P. G. 2012. 'Perception-Driven Facial Expression Synthesis', *Computers & Graphics* 36: 152–162.
- Yuki, M., Maddux, W. W. and 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: 303–311.

9. The Impact of Context on the Perception of Emotions

Ursula Hess and Shlomo Hareli

9.1 Introduction

Most human interactions are imbued with emotional exchange. In fact, it is hard to imagine a meaningful interaction in which emotion communication plays no role and even banal everyday transactions such as paying at a supermarket often involve an exchange of smiles or sometimes the expression of displeasure by one or both interaction partners. Importantly, these expressions serve as social signals that provide information about the expresser as well as about the situation (Hess, Kappas and Banse 1995) and that help to coordinate and facilitate interpersonal interaction and communication (Niedenthal and Brauer 2012; Parkinson, Fischer and Manstead 2005). This chapter explores the influence of context on the social signal value of such emotion expressions. What do they in fact tell us about the person or the situation and what influence has the context they occur in on the meaning of the exchange?

When considering the impact of context, on the one hand, and the social signals inherent in emotion expressions, on the other, the first question to ask would be what emotions actually signal. The scientific study of emotion expressions is usually traced to Darwin's seminal work *On the Expressions of the Emotions in Man and Animal* ([1872] 1965). Darwin understood emotion expressions as the visible part of an underlying emotional state, which are evolved and (at least at some point in the past) adaptive. The notion that the expressions communicate the organisms' state and thereby allow a prediction of the organisms' likely behaviour was a central point in this argument. Yet, Darwin's view has been disputed and rejected by those who considered facial expressions as exclusively or predominantly social or cultural signals, which are not linked to underlying states. Research by Ekman and colleagues (Ekman

1973; Ekman and Friesen 1971; Ekman, Friesen and Ellsworth 1972; Ekman, Sorenson and Friesen 1969) as well as Izard (Izard 1971a, 1971b) initially vindicated Darwin's idea that at least some basic emotional expressions are universal and directly associated with an underlying emotional state (see Hwang and Matsumoto, this volume, for a defence of this view, and Jack and Russell, both this volume, for criticisms). To this day, this view has been repeatedly challenged by those who consider emotion expressions to be purely social signals or social constructions unconnected to an underlying state (Barrett 2013; Fridlund 1994). Yet, at the same time emotion expressions are calorically costly which raises the question of the usefulness that is necessary to maintain such a system through mammalian history in much of which the social negotiation aspect can only be of limited relevance especially in species that do not live in social groups.

However, in some ways the question of what emotions actually express is less important when considering how they are interpreted. Specifically, as is amply demonstrated by the use of facial expressions in the arts, films and literature, people understand emotional facial expressions to express emotions and they react in the light of this understanding. Thus, people treat emotion expressions as if they express emotions and act in accordance. Yet, what are the consequences of considering emotion expressions as signs of emotions?

According to appraisal theories of emotion, emotions are elicited and differentiated through a series of appraisals of (internal or external) stimulus events based on the perceived nature of the event (e.g., Frijda 1986; Scherer 1987). Appraisal theory posits that a change in the (internal or external) environment is evaluated according to whether the event is pleasant or unpleasant (pleasantness) as well as whether the change is in line with the motivational state of the individual or obstructs the individual's goals (goal obstruction). Individuals further evaluate their ability to cope with or adjust to the change (coping potential). A further set of evaluations regards the correspondence with the relevant social and personal norms, that is, how the event is to be judged in terms of ethical, moral or social norms (norm incompatibility). The outcome of these appraisals is partially determined by the personality of the person (Scherer 1987, 1999) as well as their personal values and beliefs. Specific emotions are differentiated by the pattern of appraisal they are the result of. Thus, anger is an emotion that is characterized by appraisals of goal obstruction, high coping potential and a perception of norm violation. By contrast, sadness is characterized by

appraisals of goal obstruction, but combined with low coping potential, with norms playing less of a role. In this sense, one can say that emotions tell stories. This is what Lazarus (1991) referred to as core relational themes. In this view, sadness tells a story about loss and anger, a story about insult to the self. This idea is also relevant to the conclusions observers draw from facial expressions, that is, the inferences about a person's character, their goals and intentions, which can be drawn from observing or learning about the person's emotional reaction to an event (see what follows).

Yet, an emotion expression does not occur in a vacuum. The now defunct classic model of non-verbal communication was a straight adaptation of the Communication Model by Shannon and Weaver (1949) according to which a message is transmitted from a sender to a receiver. There is feedback from the receiver to the sender and the message may be distorted by noise. In this model, the message sent actually matches the intended message and, conversely, the understood message matches the received message. This is, however, not how human communication works. In fact, it is often the case that individuals believe to have sent a clear message when this was not the case. Humans overestimate both the intensity of their expressions (Barr and Kleck 1995) and the clarity, that is, the ease of interpretability, of their expressions (Senécal, Murard and Hess 2003) as well as the extent to which they are in fact observed by others (the spot light effect, Gilovich, Medvec and Savitsky 2000). Ambiguity in expression in turn, makes information about the situation in which the expression occurred and the naïve emotion theories of the perceiver more relevant (see the following). These naïve emotion theories also include stereotypes and cultural and social encoding rules. Thus, the application of social knowledge to what is perceived influences the perception of the message such that the same non-verbal behaviour, when shown by a different person or in a different situation, will be interpreted differently. That is, the social context (who interacts with whom and where) can exert strong influences, not only on what is sent, but also on what is perceived. In fact, it can be argued that social context becomes a determining factor in many everyday interactions. However, as noted by Darwin and others and is an implicit core of appraisal theory of emotion (Scherer 1987; cf. Frijda 1986) emotion expressions convey meaning and this meaning limits the impact of context.

In what follows, we will first discuss the process of understanding facial expressions and drawing inferences based on these expressions.

We will then discuss the elements of context, which in our conceptualization extends over current discussions of context in terms of information about the emotion eliciting situation to include the tacit information the perceiver has about relevant social rules and norms as well as the perceiver's own goals, motives and emotions. Finally, we will focus on the limits of context – that is, on the important caveat that context is not all, but that emotion expressions have in fact a true communicative value.

The present chapter focuses on facial expressions. However, much of what we discuss can be applied to emotion decoding processes in general, both those based on non-verbal cues such as postures, tone of voice and gestures and those based on second hand information such as verbal descriptions of the expresser's behaviour.

9.2 The Two-Path Model of Emotion Recognition

There are two ways to identify emotions from non-verbal cues. Most research on emotion recognition implicitly assumes a pattern-matching process, where specific features of the expression are associated with specific emotions (Buck 1984a). For example, upturned corners of the mouth or lowered brows are recognized as smiles or frowns respectively and a perceiver can thus conclude that the individual is happy or angry. In this process the perceiver is a passive decoder, who could and in fact can (e.g., Dailey et al. 2002) be replaced by an automated system, and context information plays no role or only a minimal one.

However, when the perceiver knows the expresser or is aware of the situation in which the emotion is shown, he/she can adopt an active role in the emotion identification process. Knowing about the event allows people to use their naïve emotion theories about the emotions that are typically elicited in certain events to predict the most likely emotion. For example, knowing that someone's car was vandalized typically leads to the expectation that the person will be angry. Thus, even if the person is not very expressive we can still assume that she is angry. Knowing the goals and values of others allows the perceiver to take their perspective and to infer their likely emotional state. Knowing about the temperament and emotional dispositions of the expresser further allows us to refine predictions. Thus in the case cited, we may expect more intense anger from a choleric person than from an easy-going one and more anger if the car was cherished than if it was not.

Whereas a pattern-matching approach to decoding emotion expressions works well for the intense and unambiguous expressions that are typically depicted in standardized sets of emotion expressions, such as the Pictures of Facial Affect (Ekman and Friesen 1976), it breaks down in many everyday situations where the non-verbal signal is often weak and ambiguous (Motley and Camden 1988). In this case, perspective taking can allow an observer to deduce the likely emotional reaction based on both the ambiguous expression and the context information.

But what happens if the expresser does not know the other person well or at all? In this case, any social category that the perceiver is aware of and for which expectations regarding emotional reactions exist can affect emotion identification (Kirouac and Hess 1999) in that the perceiver is more likely to attribute the more expected emotion evidenced in the ambiguous expression. For example, knowing that a (male) expresser is black or of high status leads observers to more readily label their expression as angry (Hugenberg and Bodenhausen 2003; Ratcliff et al. 2012). In the same vein, when a person is identified as a surgeon participants rate the facial expressions of the person as less intensely emotional than the same person and expressions when associated with a different identity, following the stereotype expectation that surgeons control and restrain their emotions (Hareli, David and Hess 2013).

In sum, the identification of emotions can be accomplished using either a passive pattern-matching process or a process where the perceiver actively generates a label for the likely emotional state of the sender based on both the expression and their knowledge of the context. This knowledge can take either the form of individualized knowledge about the expresser or be based on the expresser's social group and the stereotypes, expectations and beliefs associated with members of this group.

9.3 Inferences from Emotion Perception

9.3.1 A Model of the Reverse Engineering of Appraisals

As mentioned, appraisal theories of emotion posit that emotions are elicited by the spontaneous and intuitive appraisal of (internal or external) relevant stimulus events according to the perceived nature of the event (Arnold 1960; Scherer 1987). Importantly, appraisals relate to the subjective perception of the stimulus and not its objective characteristics.

Thus, the mere fact that someone reacts with an emotion to an event signals that the event is relevant to that specific person, which in turn

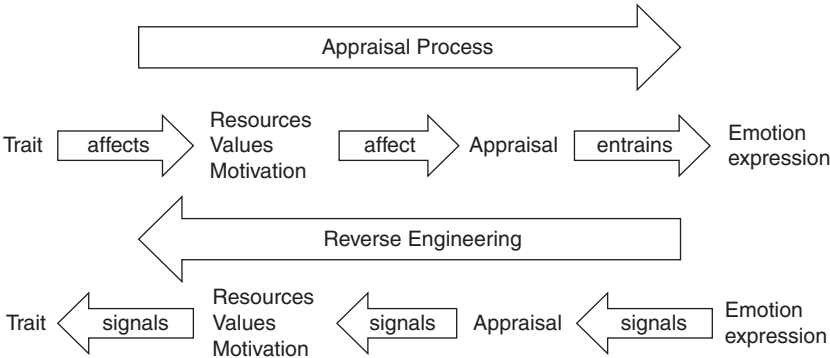


Figure 9.1. Reverse engineering of appraisals (Hess and Hareli 2015).

provides information about the person's goals and values. For example, the fact that a person reacts with anger to a perceived injustice signals that the person cares about this fact. This information is therefore encoded in the emotional expressions that are generated in this process.

Importantly, even though appraisals are typically not the product of reasoning processes, people can and do reconstruct appraisal processes consciously after the fact (Robinson and Clore 2002) and they can do so for other people's emotions as well (e.g., Roseman 1991; Scherer and Grandjean 2008). As such, emotions can be seen as encapsulated or compacted signals that tell a rather complex story about the emoter (Hareli and Hess 2010) (see Figure 9.1).

Importantly, the information provided by emotional reactions refers not only to the situation at hand but also to relatively stable characteristics of the person. Specifically, stable traits such as dominance, affiliation and competence impact the motivational goals, preferences and resources of a person. Thus, a person who is competent may be expected to have more resources to deal with potential problems than a person who is not. Likewise, an affiliative person can be expected to have affiliative goals. Conversely, seeing a person react with anger in a difficult situation suggests that this person is high in resources in this situation and likely in other situations as well.

In short, emotion displays convey, by their very nature, information not only regarding the senders' emotional states, but also information regarding their interpersonal intentions (see also Frijda and Mesquita 1994). However, the attribution of behavioural intentions also depends on the context of the interaction. Such elements as the relative status of

the interaction partners, the gender composition of the dyad and the cultural background of the interaction partners all play an important role in this regard. In fact, the social context can be expected to modify the interpretation of a specific expression, insofar as context information is used as part of the process of understanding the other, that is, permits the taking of the other's perspective in a given situation. Context can also impact on the inferences that are drawn from these expressions.

9.4 Types of Context

The importance of context to the interpretation of emotion expressions is no new discovery. The modified Brunswick model developed by Scherer (1978) in the context of person perception and which has then been applied to emotion communication includes cultural context, social relationships and situational context. In research, context has been provided by a variety of means (see also Matsumoto and Hwang 2010), such as combining facial expressive stimuli with pictures or stories describing the purported emotion elicitor (Carroll and Russell 1996; Noh and Isaacowitz 2013; Szczurek, Monin and Gross 2012). Social group information can be provided by adding information such as a Niqab or a surgical mask to the facial information (Fischer et al. 2012; Hareli, David and Hess 2013; Kret and de Gelder 2012) or by providing group labels (Thibault, Bourgeois and Hess 2006). The presence of other people also provides a context to the situation (Kafetsios and Hess 2013; Masuda et al. 2008). Alternatively, information from the face has been contrasted with information from another non-verbal channel – for example, the body (Aviezer et al. 2008). As this overview shows there is a wide range of information that can be considered context.

9.4.1 Situational Context

First, there are all those elements of the situation that are informative about the emotion elicitor. This includes factual information as well as the real world knowledge that people have and that allows them to deduce further information. For example, information that a person just competed in a game is factual information, information that players in a competition have negative interdependence such that what is good for the one must be bad for the other is real world knowledge.

Thus, when I see one competitor with a big smile, I can assume that the other one is not feeling like smiling because they lost.

Another element of the situation is the presence of others. Other individuals may directly interact with the expresser but they may also just be present at the event. In this case, they may provide information about the meaning of the situation through social referencing. Specifically, when people are confronted with complex or ambiguous situations, the emotional reactions of others can be informative about the event. This information can then be used as an input to one's own emotional reaction to and appreciation of the event. For example, in a recent study, Landmann and colleagues (2015), participants were asked to evaluate stories describing an unusual behaviour, an impolite behaviour or an immoral act. Participants also saw a picture showing another person who had reacted to these events with either anger, disgust or neutrality. In effect, the same event was rated as more immoral when the participant saw that person reacting to it with anger or disgust than with neutrality and these effects were mediated using appraisals associated with the perceived expressions. That is, participants reverse engineered the appraisals from the expressions and used these to inform their own reactions to the event.

These effects should be distinguished from the effect that the valence of the situation may have through priming or other perceptual effects. For example, when a face is shown together with a scene without there being a logical link between the two, the valence of the situation can activate affective response categories (a funny scene can activate response categories linked to positive affect) and hence influence emotion decoding. Thus, Righart and de Gelder (2008) found that when participants were asked to categorize facial expressions that were shown against the backdrop of an emotional scene while ignoring the scene, the categorizations were biased by the emotional content of the scenes. Somewhat similar effects occur when faces are shown within a group of other individuals especially when the presence of the others is not explained; these effects tend to be stronger for people high in interdependence (Hess, Blaison and Kafetsios forthcoming; Masuda et al. 2008).

9.4.2 The Perceiver as Context

Another important element of context is the perceiver. The two-path model of emotion perception considers the perceiver not as a passive read-out module but as taking an active part in the perception process.

As such, not only the real world knowledge mentioned, but the stereotypes the perceiver holds, the norms the perceiver is aware of and the perceiver's own goals and motives are all relevant for this process. We will discuss these in turn.

Stereotypes, expectations and social norms. Stereotypes are not the same as social norms. However, in this context we consider mostly prescriptive stereotypes that imply a behavioural norm. Thus, if someone holds the stereotype of women as more irrationally emotional and men as more controlled, they should expect men to act with more emotional restraint (Shields 2005).

Clearly stereotype and norm knowledge require some level of situational knowledge, which anchors the relevant norm. However, insofar as the norm relates to specific social groups, which can be identified based on their face alone (e.g., racial groups, the elderly, men and women) this situational knowledge may be activated by the very expression that is to be decoded. Thus, with regard to the quoted example, if the perceiver knows that the person whose car was vandalized is a woman, norms of behaviour relevant to women and anger will become more accessible. These norms may then influence the identification of an emotional cue associated with the target person. Specifically, women were expected to show sadness rather than anger in such a situation unless they were explicitly described as very dominant (Hess, Adams and Kleck 2005). As the identification of emotions includes not only the labeling of a 'central' emotion but also of secondary emotions, these can be influenced by stereotype expectations as well. For example, Algoe, Buswell and DeLamater (2000) have shown that observers perceived fear expressions as reflecting more intense anger and contempt when targets were described as the boss (i.e., high status) rather than as employees (i.e., low status) in line with stereotype expectations related to power.

Cultural display rules. A specific case of norms are cultural display rules, that is, the social rules that guide the appropriate display of emotion expressions (Ekman and Friesen 1971). These differences can in part be related to differences in cultural values such as individualism and collectivism (Matsumoto, Seung Hee and Fontaine 2008) but also openness to change (Koopmann-Holm and Matsumoto 2011) or masculinity (Sarid 2015), among others. Mostly, however, we can assume that cultural display rules are not linked to one specific cultural value but are the result of more complex processes involving more than one cultural attribute. Importantly in this context, social display rules have

a converse side in social decoding rules (Buck 1984b) such that perceivers tend to be less good at decoding expressions that are proscribed in a given culture. For example, in a recent study, Hess, Blaison and Kafetsios (forthcoming) found that Greek participants rated sad faces as more intensely sad and German participants rated anger faces as more angry than vice versa in line with cultural display rules that respectively endorse the expressions of these emotions in these two countries.

The perceivers' goals, needs and own emotional state in emotion identification. A second perceiver-related context factor is the perceivers' goals, and needs, and even their own emotional state (Showers and Cantor 1985). These factors specifically affect what is extracted from the available bottom-up information, for example, by determining the degree of effort that the perceiver invests. Thus, being highly motivated and having the ability to do so, a perceiver may pay more attention to the available cues. By contrast, if motivation and/or ability are low, less attention may be paid. Thus, Thibault, Bourgeois and Hess (2006) found that perceivers who strongly identified with members of a group were better at labeling emotion expressions from members of that group. This finding fits well the more general idea that people often invest relatively little effort in learning about the characteristics of out-group others (Park and Rothbart 1982). In a similar vein, research on gender differences in emotion recognition shows that motivational factors may have a substantial impact on recognition accuracy and may explain why in some studies women outperform men in this task (Ickes and Simpson 2004).

However, motives do not only affect attention, but can also reframe the perception of emotion expressions. For example, the smile of another person is usually perceived positively as happiness. But, when, as in the example, perceivers know the other person to be in competition with them, the perceivers' own goal to achieve success becomes salient, and the same smile may become a smirk and the happiness becomes glee in their mind.

Not only attention and framing of the available cues but also their processing can be influenced by perceiver characteristics. A good example would be the way the emotional state of the observer affects decoding. An individual's emotional state influences how social information is processed (e.g., Bower and Forgas 2000, 2001). Specifically, according to Forgas's 'affect infusion model' (1995), perceivers' information processing strategies differ in the extent to which a full search of information occurs and how open or closed this search is and hence in the use of perceiver knowledge. At one extreme of this process, the

perceiver may directly and automatically retrieve a preexisting label when encountering a stimulus. At the other extreme, the perceiver may engage in an extensive and open search of information. This latter strategy involves substantive processing using preexisting knowledge in a relatively unbiased manner (Bower and Forgas 2000). Thus, the degree to which subtle cues and situational information are integrated when, for example, trying to label a smile, depends also on the emotions felt by the perceiver.

To render the issue more complex, the perceiver's emotional state may in fact be partially determined by an automatic reaction to the emotion of the expresser, such as when emotional contagion occurs (cf. Hatfield, Cacioppo and Rapson 1994). This process could conceivably precede the conscious identification of the other's emotional state. For example, the fear of a target may cause fear in the perceiver. This fear, in turn, may affect the identification process by causing the perceiver to be less attentive to the details of the target's emotion – or give preferential status to the elements of the expression consistent with the perceiver's own fear. Thus, even if the perceiver correctly identifies the emotion as fear, the fear may be attributed to the wrong cause or associated with the wrong object or be perceived as more intense than it actually is.

Obviously, both processes mentioned may operate at the same time such that low motivation and/or ability will both result in partial attention to cues and in the limited processing of these cues, such that more easily accessible stereotypical knowledge may serve as ready-made templates for recognition based on a superficial observed feature (Chaiken 1980; see also Forgas 1995).

Finally, it should be noted that individual differences in personality also can affect decoding and that these also provide an element of context. For example, traits such as hostility (Larkin, Martin and McClain 2002) and aggression (Hall 2006) can bias emotion perception.

In sum, context can be defined in a variety of ways and includes both the situation and the perceiver. The perceiver's knowledge, naïve emotion theories, motivations goals and emotions all enter into the active process described in the two-path model of emotion recognition. However, this raises the question regarding the limits of the influence.

9.5 Limits to the Malleability of Emotion Perception

The pervasive influence of context on emotion perception can give the impression that early critics of Darwin were right in saying that facial

expressions per se are meaningless. That at best they can – as claimed by Bruner and Tagiuri (1954) – be culturally learned signals, which are not meaningfully linked to an underlying state or are to be considered as constructed within the moment either at the interface between individual and environment (Mesquita and Boiger 2014) or in the head of the individual (Barrett 2009, 2013), that is, in the vein of strong psychological constructivism (Faucher 2013) created in a ‘simulator’ which constructs ‘on-the-fly’ emotion concepts adapted to particular instances of a category.

We think that this impression is false. It is important to note that even though context frames the way people interpret cues and the attention that is paid to the cues as well as the level of processing that is applied to this endeavor, context is also confronted and limited by the story that emotions tell. More specifically, context is limited by a framework based on the core appraisals that distinguish one emotion from another and that create the emotion’s story. That is, only within the frame provided by that story and only within the limits of these appraisals can context change our perception of emotions. When, in fact, context and expression diverge past this frame, then the expresser will be considered ‘strange’ or ‘deviant’ (Szcurek, Monin and Gross 2012) and the expression correspondingly discounted. In fact, this notion can be supported by research originally designed to underline the power of context. Thus, Aviezer et al. (2008), in an attempt to show the malleability of emotion perception, created stimuli which combined an emotion expression with a body stance. In study 1, a disgust face was combined with stances communicating disgust, anger, sadness and fear. There are two types of disgust, physical disgust in reaction to noxious stimuli, and moral disgust in reaction to morally inappropriate behaviour (Rozin et al. 1999), and this disgust resembles anger in that it is associated with goal obstruction and high coping potential combined with an appraisal of norm violation. By contrast, fear and sadness are both emotions that are characterized by low coping potential and norm appraisal is not very relevant. The findings show that the disgust face combined with an aggressive body posture was indeed overwhelmingly miscategorized as anger (87 per cent). However, when the disgust face was combined with fear (13 per cent) and sadness postures (29 per cent), which are much less compatible with the appraisal pattern for moral disgust, it was miscategorized to a substantially smaller degree.

Similarly, the inferences drawn from an emotion are influenced by situational information only to the degree that this information is

relevant within the appraisal pattern. Hareli, Elkabetz and Hess (2015) showed participants a series of images from a fictional ball game. At the end of the game one team member's facial expression is seen and this person is identified by the colour of their shirt as either a member of the team that played the last turn or as an opponent of the team. The team member shows either a happy, neutral or awe expression. As mentioned, participants' real world knowledge tells them that there is negative interdependence between the two teams. Participants were asked about the performance of the last player on the field. When the expresser showed happiness, their judgement depended on the presumed motivation of the expresser – that is, on whether they were on the same team as the player or on the opposing team. Happiness is an emotion for which the central appraisals are pleasantness and goal conduciveness – essentially the notion that something nice and in line with the plans of the expresser had happened. Hence the player's performance was rated as good when a supporter of the same team was seen smiling and as bad when a supporter of the opposing team was smiling. Yet, when awe was shown, something very different happened. Awe is associated with appraisals of something that exceeds norms. It does not depend on goal conduciveness. In fact, people report being in awe of forces of nature such as storms, earthquakes or volcanic eruptions, none of which is goal conducive. In line with this notion, participants' judgement did not at all depend on who showed the expression, whether supporter or opponent showed awe the performance of the player was rated as excellent.

In sum, context plays a very important role in emotion perception; however, it plays this role within the – admittedly large – framework of the core appraisals characterizing this emotion.

9.6 A Model of Social Signals in Context

On the basis of the concepts discussed, we formulated a model of the meaning of emotion expressions in context (MEEC, see [Figure 9.2](#)). In this model, expressions are perceived within a situational context (the real world) and then interpreted within an interpreted context (the perceived world). The information from the real world will determine the encapsulated meaning of the expression (the story that the emotion tells) and this process will be influenced by the perceiver as context processes outlined. If the information provided by the context and the interpretation of the expression fall within the frame of the core

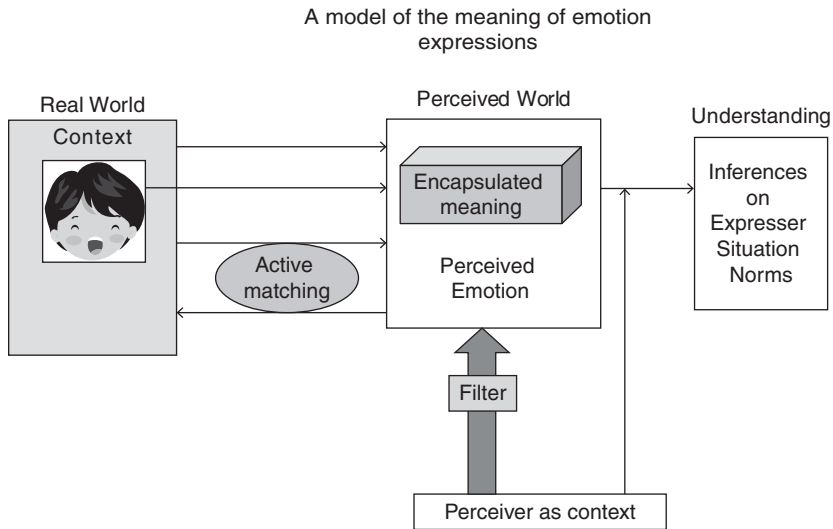


Figure 9.2. A model of the meaning of emotion expressions.

appraisals associated with the emotion the process can go on to allow for inferences to be drawn from the expression. In case of a mismatch the perceiver has to reevaluate the match explicitly. One outcome of this process can be to discount the expression as ‘deviant’ and to ignore its input. Another outcome could be to reevaluate the situation. For example, most people react positively to puppies. If a person shows fear in response to the puppy one might consider the possibility that the person suffers from an extreme form of fear of dogs. However, just as in horror movies it might be that just behind the puppy a large aggressive drooling, likely rabid, dog can be seen which changes the situation completely.

The model distinguishes between two types of context, the ‘real world’ or situational context and the context provided by the perceiver’s knowledge, goals and emotions; but the model does not distinguish within these categories. In fact, as the short overview shows, there are many operationalizations of context, without a clear definition of what is or is not to be considered context. For example, from a multi-model emotion decoding perspective (Bänziger, Grandjean and Scherer 2009) facial expressions and body postures are both emotion signals and not a context for each other. Also, when an unattached head is floating above an image without an obvious link between the two, it is not clear whether the scene in the image is a context for

the face, even though it certainly influences the perception of the face (Righart and De Gelder 2008). Discussions about the role of context for the construction of emotional meaning, however, require a clearer definition of both what is considered to be signal and what is considered to be ancillary information, as not everything that is perceived at the same time as the expression has the same epistemological standing with regard to the meaning of an expression. Thus, as noted earlier, a scene that is depicted together with a face can – through affective priming – activate response categories, which in turn facilitate or hinder the categorization of the expression. This bottom-up process is to be distinguished from the type of top-down process that is engaged when participants use real world information to deduce expressers' motives when drawing inferences from their expressions. Future research and theorizing needs to pay more attention to the specific processes engaged in the construction of the meaning of emotion expressions and in the limits of this process. In this vein, it would be important to not only show when a specific context influences perception but also when it does not.

Another, more difficult to study aspect regards the question of what is in fact constructed. In this chapter we have focused exclusively on the *perception* of (facial) emotion expressions. Yet, the same general line of argument has been applied also to the elicitation process (Barrett 2006). Generally speaking, encoding is not the same as decoding and clearly the same processes need not be at work. For example, at least in Western cultures, the presence of wrinkles around the eyes when smiling (the so-called Duchenne smile) is understood as a signal of an authentic or 'true' smile of enjoyment (Thibault et al. 2012), but there is good evidence that this smile is regularly employed by expressers as a social smile in situations devoid of enjoyment (Hess and Bourgeois 2010; Krumhuber and Manstead 2009). Thus, the attribution of enjoyment by perceivers is not matched by the actual elicitation conditions. This and other examples suggest that care needs to be taken when translating from one domain to the other.

In sum, the model of the meaning of emotion expressions in context (MEEC) proposes a pertinent but not exclusive role for context information by proposing core appraisals as the limiting frame of interpretation. The model, just as do social constructivist accounts, considers perceivers as active participants in the emotion decoding process, but as ones who are constrained with regard to their constructive freedom.

References

- Algoe, S. B., Buswell, B. N. and Delamater, J. D. 2000. 'Gender and Job Status as Contextual Cues for the Interpretation of Facial Expression of Emotion', *Sex Roles* 42: 183–208.
- Arnold, M. B. 1960. *Emotion and Personality*. New York: Columbia University Press.
- Aviezer, H., Hassin, R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M. and Bentin, S. 2008. 'Angry, Disgusted, or Afraid? Studies on the Malleability of Emotion Perception', *Psychological Science* 19: 724–732.
- Bänziger, T., Grandjean, D. and Scherer, K. R. 2009. 'Emotion Recognition from Expressions in Face, Voice, and Body: The Multimodal Emotion Recognition Test (MERT)', *Emotion* 9: 691–704.
- Barr, C. L. and Kleck, R. E. 1995. 'Self-Other Perception of the Intensity of Facial Expressions of Emotion – Do We Know What We Show?', *Journal of Personality and Social Psychology* 68: 608–618.
- Barrett, L. F. 2006. 'Solving the Emotion Paradox: Categorization and the Experience of Emotion', *Personality and Social Psychology Review* 10: 20–46.
2009. 'Variety Is the Spice of Life: A Psychological Construction Approach to Understanding Variability in Emotion', *Cognition and Emotion* 23: 1284–1306.
2013. 'Psychological Construction: The Darwinian Approach to the Science of Emotion', *Emotion Review* 5: 379–389.
- Bower, G. H. and Forgas, J. P. 2000. 'Affect, Memory, and Social Cognition', in Eich, Kihlstrom, Bower, Forgas and Niedenthal (eds.) *Cognition and Emotion*. New York: Oxford University Press.
2001. 'Mood and Social Memory. Handbook of Affect and Social Cognition', in Forgas (ed.) *Handbook of Affect and Social Cognition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bruner, J. S. and Tagiuri, R. 1954. 'The Perception of People', in Lindzey (ed.) *Handbook of Social Psychology*. Cambridge, MA: Addison-Wesley Publishing.
- Buck, R. 1984a. *The Communication of Emotion*. New York: Guilford Press.
- 1984b. 'Nonverbal Receiving Ability', in Buck (ed.) *The Communication of Emotion*. New York: Guilford Press.
- Carroll, J. M. and Russell, J. A. 1996. 'Do Facial Expressions Signal Specific Emotions? Judging Emotion from the Face in Context', *Journal of Personality and Social Psychology* 70: 205–218.
- Chaiken, S. 1980. 'Heuristic versus Systematic Information Processing and the Use of Source versus Message Cues in Persuasion', *Journal of Personality and Social Psychology* 39: 752–766.
- Dailey, M. N., Cottrell, G. W., Padgett, C. and Adolphs, R. 2002. 'EMPATH: A Neural Network That Categorizes Facial Expressions', *Journal of Cognitive Neuroscience* 14: 1158–1173.
- Darwin, C. [1872] 1965. *The Expression of the Emotions in Man and Animals*. Chicago: University of Chicago Press.
- Ekman, P. 1973. *Darwin and Facial Expression: A Century of Research in Review*. New York: Academic Press.

- Ekman, P. and Friesen, W. V. 1971. 'Constants across Cultures in the Face and Emotion', *Journal of Personality and Social Psychology* 17: 124–129.
1976. *Pictures of Facial Affect*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., Friesen, W. V. and Ellsworth, P. 1972. *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. New York: Pergamon Press.
- Ekman, P., Sorenson, E. R. and Friesen, W. V. 1969. 'Pan-Cultural Elements in Facial Displays of Emotion', *Science* 164: 86–88.
- Faucher, L. 2013. 'Comment: Constructionisms?', *Emotion Review* 5: 374–378.
- Fischer, A. H., Gillebaart, M., Rotteveel, M., Becker, D. and Vliek, M. 2012. 'Veiled Emotions: The Effect of Covered Faces on Emotion Perception and Attitudes', *Social Psychological and Personality Science* 3: 266–273.
- Forgas, J. P. 1995. 'Mood and Judgment: The Affect Infusion Model (AIM)', *Psychological Bulletin* 117: 39–66.
- Fridlund, A. J. 1994. *Human Facial Expression: An Evolutionary View*. New York: Academic Press.
- Frijda, N. H. 1986. *The Emotions*. Cambridge: Cambridge University Press.
- Frijda, N. H. and Mesquita, B. 1994. 'The Social Roles and Functions of Emotions', in Kitayama, S. and Markus, H. R. (eds.) *Emotion and Culture: Empirical Studies of Mutual Influence*. Washington, DC: American Psychological Association.
- Gilovich, T., Medvec, V. H. and Savitsky, K. 2000. 'The Spotlight Effect in Social Judgment: An Egocentric Bias in Estimates of the Salience of One's Own Actions and Appearance', *Journal of Personality and Social Psychology* 78: 211–222.
- Hall, C. W. 2006. 'Self-Reported Aggression and the Perception of Anger in Facial Expression Photos', *Journal of Psychology: Interdisciplinary and Applied* 140: 255–267.
- Hareli, S., David, S. and Hess, U. 2013. 'Competent and Warm but Unemotional: The Influence of Occupational Stereotypes on the Attribution of Emotions', *Journal of Nonverbal Behavior* 37: 307–317.
- Hareli, S. and Hess, U. 2010. 'What Emotional Reactions Can Tell us about the Nature of Others: An Appraisal Perspective on Person Perception', *Cognition and Emotion* 24: 128–140.
- Hareli, S., Elkabetz, S. and Hess, U. 2015. 'The Role of Context in Social Inferences Drawn from Emotions as a Function of the Situative Informativeness of the Emotion – The Case of Awe and Happiness', *Manuscript*.
- Hatfield, E., Cacioppo, J. T. and Rapson, R. L. 1994. *Emotional Contagion*. Madison, WI: C.W. Brown.
- Hess, U. and Bourgeois, P. 2010. 'You Smile – I Smile: Emotion Expression in Social Interaction', *Biological Psychology* 84: 514–520.
- Hess, U. and Hareli, S. 2015. 'The Role of Social Context for the Interpretation of Emotional Facial Expressions', *Understanding Facial Expressions in Communication*, pp. 119–141. India: Springer.
- Hess, U., Adams, R. B., Jr. and Kleck, R. E. 2005. 'Who May Frown and Who Should Smile? Dominance, Affiliation, and the Display of Happiness and Anger', *Cognition and Emotion* 19: 515–536.

- Hess, U., Blaison, C. and Kafetsios, K. Forthcoming. 'Judging Facial Emotion Expressions in Context: The Influence of Culture and Self-Construal Orientation', *Journal of Nonverbal Behavior*.
- Hess, U. and Bourgeois, P. 2010. 'You Smile – I Smile: Emotion Expression in Social Interaction', *Biological Psychology* 84: 514–520.
- Hess, U., Kappas, A. and Banse, R. 1995. 'The Intensity of Facial Expressions Is Determined by Underlying Affective State and Social Situation', *Journal of Personality and Social Psychology* 69: 280–288.
- Hugenberg, K. and Bodenhausen, G. V. 2003. 'Facing Prejudice: Implicit Prejudice and the Perception of Facial Threat', *Psychological Science* 14: 640–643.
- Ickes, W. and Simpson, J. A. 2004. 'Motivational Aspects of Empathic Accuracy', in Brewer and Hewstone (eds.) *Emotion and Motivation. Perspectives on Social Psychology*. Malden, MA: Blackwell Publishing.
- Izard, C. E. 1971a. *The Face of Emotion*. New York: Appleton-Century-Crofts.
- 1971b. *Human Emotions*. New York: Plenum.
- Kafetsios, K. and Hess, U. 2013. 'Effects of Activated and Dispositional Self-Construal on Emotion Decoding Accuracy', *Journal of Nonverbal Behavior* 37: 191–205.
- Kirouac, G. and Hess, U. 1999. 'Group Membership and the Decoding of Nonverbal Behavior', in Philippot, Feldman and Coats (eds.) *The Social Context of Nonverbal Behavior*. Cambridge: Cambridge University Press.
- Koopmann-Holm, B. and Matsumoto, D. 2011. 'Values and Display Rules for Specific Emotions'. *Journal of Cross-Cultural Psychology* 42: 355–371.
- Kret, M. E. and De Gelder, B. 2012. 'Islamic Headdress Influences How Emotion Is Recognized from the Eyes', *Frontiers in Psychology* 3: 110.
- Krumhuber, E. and Manstead, A. S. R. 2009. 'Can Duchenne Smiles Be Feigned? New Evidence on Felt and False Smiles', *Emotion* 9: 807–820.
- Landmann, H., David, S., Hareli, S. and Hess, U. 2015. 'How I See You Depends on What You See and Vice Versa: The Bidirectional Relation of Emotion Perception and Morality', *Manuscript*.
- Larkin, K. T., Martin, R. R. and McClain, S. E. 2002. 'Cynical Hostility and the Accuracy of Decoding Facial Expressions of Emotions', *Journal of Behavioral Medicine* 25: 285–292.
- Lazarus, R. S. 1991. *Emotion and Adaptation*. New York: Oxford University Press.
- Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S. and Van De Veerdonk, E. 2008. 'Placing the Face in Context: Cultural Differences in the Perception of Facial Emotion', *Journal of Personality and Social Psychology* 94: 365–381.
- Matsumoto, D. and Hwang, H. S. 2010. 'Judging Faces in Context', *Social and Personality Psychology Compass* 4: 393–402.
- Matsumoto, D., Seung Hee, Y. and 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: 55–74.
- Mesquita, B. and Boiger, M. 2014. 'Emotions in Context: A Sociodynamic Model of Emotions', *Emotion Review* 6: 298–302.

- Motley, M. T. and Camden, C. T. 1988. 'Facial Expression of Emotion: A Comparison of Posed Expressions versus Spontaneous Expressions in an Interpersonal Communications Setting', *Western Journal of Speech Communication* 52: 1–22.
- Niedenthal, P. M. and Brauer, M. 2012. 'Social Functionality of Human Emotion', *Annual Review of Psychology* 63: 259–285.
- Noh, S. R. and Isaacowitz, D. M. 2013. 'Emotional Faces in Context: Age Differences in Recognition Accuracy and Scanning Patterns', *Emotion* 13: 238–249.
- Park, B. and Rothbart, M. 1982. 'Perception of Out-Group Homogeneity and Levels of Social Categorization: Memory for the Subordinate Attributes of In-Group and Out-Group Members', *Journal of Personality and Social Psychology* 42: 1051–1068.
- Parkinson, B., Fischer, A. H. and Manstead, A. S. R. 2005. *Emotion in Social Relations: Cultural, Group, and Interpersonal Processes*. New York: Psychology Press.
- Ratcliff, N. J., Franklin, R. G., Nelson, A. J. and Vescio, T. K. 2012. 'The Scorn of Status: A Bias toward Perceiving Anger on High-Status Faces', *Social Cognition* 30: 631–642.
- Righart, R. and de Gelder, B. 2008. 'Recognition of Facial Expressions Is Influenced by Emotional Scene Gist', *Cognitive, Affective, & Behavioral Neuroscience* 8: 264–272.
- Robinson, M. and Clore, G. 2002. 'Belief and Feeling: Evidence for an Accessibility Model of Emotional Self-Report', *Psychological Bulletin* 128: 934–960.
- Roseman, I. J. 1991. 'Appraisal Determinants of Discrete Emotions', *Cognition and Emotion* 5: 161–200.
- Rozin, P., Lowery, L., Imada, S. and Haidt, J. 1999. 'The CAD Triad Hypothesis: A Mapping between Three Moral Emotions (Contempt, Anger, Disgust) and Three Moral Codes (Community, Autonomy, Divinity)', *Journal of Personality and Social Psychology* 76: 574–586.
- Sarid, O. 2015. 'Assessment of Anger Terms in Hebrew: A Gender Comparison', *The Journal of Psychology* 149: 303–324.
- Scherer, K. R. 1978. 'Personality Inference from Voice Quality: The Loud Voice of Extraversion', *European Journal of Social Psychology* 8: 467–487.
1987. 'Towards a Dynamic Theory of Emotion: The Component Process Model of Affective States', *Geneva Studies in Emotion and Communication*. www.affective-sciences.org/node/402.
1999. 'Appraisal Theory', in T. Dalgleish and M. J. Power (eds.) *Handbook of Cognition and Emotion*. Chichester, UK: John Wiley & Sons Ltd, pp. 637–663.
- Scherer, K. R. and Grandjean, D. 2008. 'Facial Expressions Allow Inference of Both Emotions and Their Components', *Cognition and Emotion* 22: 789–801.
- Senécal, S., Murard, N. and Hess, U. 2003. 'Do You Know What I Feel? Partners' Predictions and Judgments of Each Other's Emotional Reactions to Emotion-Eliciting Situations', *Sex Roles* 48: 21–37.
- Shannon, C. E. and Weaver, W. 1949. *A Mathematical Model of Communication*. Urbana: University of Illinois Press.

- Shields, S. A. 2005. 'The Politics of Emotion in Everyday Life: "Appropriate" Emotion and Claims on Identity', *Review of General Psychology* 9: 3–15.
- Showers, C. and Cantor, N. 1985. 'Social Cognition: A Look at Motivated Strategies', *Annual Review of Psychology* 36: 275–305.
- Szczurek, L., Monin, B. and Gross, J. J. 2012. 'The Stranger Effect: The Rejection of Affective Deviants', *Psychological Science* 23: 1105–1111.
- Thibault, P., Bourgeois, P. and Hess, U. 2006. 'The Effect of Group-Identification on Emotion Recognition: The Case of Cats and Basketball Players', *Journal of Experimental Social Psychology* 42: 676–683.
- Thibault, P., Levesque, M., Gosselin, P. and Hess, U. 2012. 'The Duchenne Marker Is NOT a Universal Signal of Smile Authenticity: But It Can Be Learned!', *Special Issue of Social Psychology on Culture as Process* 43: 215–221.

10. The Truth Is Written All Over Your Face! Involuntary Aspects of Emotional Facial Expressions

Alysha Baker, Pamela J. Black and Stephen Porter

10.1 Introduction

The idea that private, covert information is subtly communicated in human facial expressions is popular in both the academic literature (e.g., the ‘micro-expression’) and the media (e.g., the television show ‘Lie to Me’). Yet, in 2005, when we began to scour the literature for evidence of involuntary aspects of emotional expression – and the contribution to identifying insincere emotional expressions – we found that it was surprisingly scarce. With the exception of genuine and false smiles (e.g., Ekman, Davidson and Friesen 1990; Frank, Ekman and Friesen 1993), little research had addressed even the most basic issues around involuntary aspects of emotional expression, including whether micro-expressions exist. In the decade since this realization, our research team has conducted the most comprehensive investigation of the manner in which emotional information is involuntarily communicated on the human face (e.g., Porter and ten Brinke 2008; ten Brinke and Porter 2012; ten Brinke et al. 2013), sometimes to the expresser’s dismay. With the goal of solving some of the mysteries and resolving some of the controversies in this area, we have manually coded millions of frames of videotaped faces for both the presence of full expressions and specific, individual muscle movements. These tapes include expressions generated in both highly controlled, lab-based experiments and naturalistic ‘real-life’ emotional scenarios. Not only does this body of research represent our basic interest in this aspect of human communication, but also it was motivated by the practical implications of the results. Specifically, in observing the deleterious consequences that emotional deception can have in society, we hoped to learn how science might inform the practice of identifying insincere emotions. For example, in

one study we found that psychopathic offenders are such good emotional actors that they successfully talk and cry their way out of parole hearings at a rate two and a half times higher than their more meritorious counterparts (Porter, ten Brinke and Wilson 2009). When we followed up on their performance in the community, they typically re-offended quickly and often in a heinous, violent fashion. The goal of this chapter is to explore the results of our and others' research on the manifestations of voluntary and involuntary aspects of emotional expressions in everyday life and the manner in which they relate to the differentiation of sincere and insincere expressions.

10.2 The Status and Role of the Face in Interpersonal Communication

A well-known saying is that 'the eyes are the windows into the soul'. In fact, research shows that although the eyes are involved in facial expressions, an observer must scrutinize the face as a whole to obtain information about a person's state and trait characteristics (Abdel-Rahman 2011; Amado, Yildirim and Iyilikçi 2011; Amihai, Deouell and Bentin 2011; Arsalidou, Morris and Taylor 2011). Indeed, in everyday life, observers read the faces of intimates and strangers to make inferences about age, gender, race, emotions and intentions (Amado, Yildirim and Iyilikçi 2011), and adopt expressions themselves to communicate their genuine or false feelings. Evidence that the face is an integral aspect of emotional communication can be found when looking to research on communication with infants. For example, beginning in early infancy humans exhibit a visual preference for face-like patterns and attend to face-like stimuli at the expense of other stimuli (e.g., Valenza et al. 1996). In addition to attending to the human face, infants' own facial musculature is formed and fully functional from birth as well (Ekman and Oster 1979; Goldfield 1995). Infants are capable of making facial expressions (i.e., a smile accompanied by 'crow's feet' around the eyes representative of happiness) that are similar to the expressions displayed by adults to express their wants and needs before language capability (Ekman and Oster 1979). Then, throughout life, the face continues to be the focal point in social situations, directing both conversation and decisions about further interaction (e.g., approach versus avoidance behaviour). This is facilitated by a number of dedicated brain regions that process faces. For example, the fusiform face area and the occipital facial area are two regions

involved in the facial expression recognition network (Fox et al. 2011; Kanwisher, McDermott and Chun 1997; Zhang et al. 2012). Complex networks in the human brain allow for the instantaneous and sub-conscious evaluation of personality characteristics, emotional states and the degree of attractiveness and trustworthiness upon viewing a human face (Amihai, Deouell and Bentin 2011; Arsalidou, Morris and Taylor 2011; Porter and ten Brinke 2009).

These initial impressions derived from the face are long lasting (e.g., Willis and Todorov 2006), a process that highlights the evolutionary function of assessing the faces of others for information relevant to survival. The manner in which the observer acts during subsequent social interactions with the target is a direct reflection of his/her first impression, largely formed from information derived from the target's face. Observers assess relatively *stable* traits, such as trustworthiness, using facial features involved in emotional expressions, including the shape of the mouth and the position of the eyebrows (up- vs. downturned; Oosterhof and Todorov 2008). If a target is perceived as seeming angry (e.g., has a naturally, morphologically downturned mouth and a furrowed brow), he/she may be perceived as untrustworthy or a threat and avoided to enhance the observer's safety (Adams et al. 2006). However, despite the frequency with which observers scrutinize targets' faces to infer their traits, research suggests that stable facial features typically are not predictive of actual personality characteristics or behaviour (see Bond, Berry and Omar 1994; Zebrowitz et al. 1996). One exception is facial dominance or masculinity, which approximates physical strength and the capacity to carry out malicious intentions (e.g., Borkenau and Liebler 1992; ten Brinke and Porter 2012). For example, males with greater facial width (associated with higher levels of testosterone) are more likely to cheat and aggress against others (Stirrat and Perrett 2010). Observers are able to differentiate the faces of Nobel Peace Prize winners and violent criminals at only slightly above the level of chance (Porter et al. 2008). Despite our limited skills at making trait inferences from the *static* faces of others, attending to the *dynamic* facial expressions of others in order to infer their emotional states can be highly adaptive and informs the observer's subsequent behaviour. For example, if a target is displaying a genuine expression of anger, an observer might avoid that person to evade possible harm. Similarly, an expression of fear may signal danger in the vicinity, which also could result in an observer fleeing the situation to avoid harm.

10.3 Emotional Expressions: Universal, Learned...or Both?

Emotions have been described as discrete and automatic affective responses to a personal event (Ekman and Cordaro 2011). Beginning with Darwin (1872) and popularized by Ekman and colleagues in the 1960s, basic emotion theories propose that certain emotion categories are biologically inherited and cannot be further reduced to psychological elements. Further, it was hypothesized that such basic emotions are expressed in facial expressions and that these expressions are formed using both voluntarily and involuntarily activated muscles in the eyes, forehead, nose region, mouth and neck (Ekman 1992a). Ekman and others argued that emotional facial expressions associated with the basic emotions have cross-cultural (and arguably cross-species) reliability and a well-established biological basis (e.g., Ekman 1993, 2003; Hamann 2012; Sauter et al. 2009; Scherer and Ekman 1984; Sievers et al. 2013; Susskind et al. 2008; Waller and Micheletta 2013; see also Hwang and Matsumoto, this volume).

The six basic emotions that have been most consistently classified as universal by Ekman and colleagues include happiness, sadness, anger, fear, surprise and disgust, with partial evidence for a seventh emotion – contempt (Ekman and Friesen 1971; Matsumoto 1992; Rosenberg and Ekman 1995). These expressions are considered universal because support for their existence has been found in varying countries, including Brazil, Argentina, Japan and the United States (Ekman and Friesen 1971; Ekman, Friesen and Ellsworth 1972; Ekman, Sorenson and Friesen 1969; Izard 1994). For example, all humans engage the same facial muscles to express genuine happiness, and this unique pattern of contracted facial muscles is reciprocally interpreted as happiness. In one of the first attempts to find support for theorized universal facial expressions, Ekman and Friesen (1971) recorded the facial expressions of the people of Papua New Guinea, a group of individuals who had never had contact with the Western world, including not having access to television, popular media or books that depicted people from North America or Europe. These individuals were read stories aimed at evoking discrete emotions and the resulting expressions were captured to examine whether these individuals displayed the proposed basic facial expressions. When Ekman and Friesen (1971) later asked American participants to identify the facial expressions displayed on the faces of the participants from New Guinea, they did so correctly. This provided further evidence that the facial expressions for discrete

emotions are universally expressed and interpreted, suggesting they are an innate human feature as opposed to socially constructed (Ekman and Friesen 1971).

Dimensional theories of emotion, however, conceptualize emotions as emerging from more elemental dimensions such as arousal and valence (e.g., Kuppens et al. 2013), or pleasure and activation (Barrett and Russell 1998) and are associated with doubt or even denial that universally recognized emotional expressions exist (Barrett 1998; Barrett, Lindquist and Gendron 2007; Jack, Garrod, Yu, Caldara and Schyns 2012; Nelson and Russell 2013; Russell 1994, 2003). For example, Kayyal and Russell (2013) followed similar procedures to that of Ekman and Friesen (1971) by showing spontaneous facial expressions exhibited by aborigines in Papua New Guinea (Ekman 1980) to different cultural groups (Americans and Palestinians) and found that there was only moderate reliability of labeling emotions for many of the faces for basic emotions. Further, the observers typically endorsed more than one emotion (or a blend of emotions) for each face, suggesting that a single label associated with the hypothesized universal emotions was insufficient to capture the complexity of emotional expressions.

Another line of relevant research in this debate has focused on the facial expressions of congenitally blind individuals who have never had the opportunity to observe the facial expressions of others. For example, Ortega et al. (1983) studied emotional expressions (using Ekman's Facial Action Coding System [FACS; Ekman and Friesen 1978]) in twenty-two congenitally blind children and a comparison group of sighted children aged 7–13 years. They found that spontaneous expressions (those elicited naturally or involuntarily to stimuli) were similar in the two groups (supporting the theory of basic, universal emotions). However, posed or voluntary expressions were much less discrete in the blind participants. The authors argued that this pattern supported a dual mechanism operating in the development of facial expression of emotions (potentially supporting the theory of dimensional emotional expressions). They saw one mechanism as genetically determined, which explains the similarities in the involuntary facial expressions among blind and sighted people, and the other as based on learning how to voluntarily and clearly portray emotions.

As such, we are well aware of the controversies over the notion of basic emotions and universally recognized emotional expressions. Our disclaimer for the remainder of the chapter is that much of our work on genuine and falsified emotional expressions has relied upon and

strongly supports the idea that there are distinct emotional expressions communicated involuntarily and spontaneously by individuals experiencing the type of emotional states that we have attempted to induce (e.g., Porter and ten Brinke 2008; Porter, ten Brinke and Wallace 2012) or would expect them to experience ‘naturally’ given a particular emotional context (ten Brinke, Porter and Baker 2012). Further, as described later on, we find that knowledge of these ‘anticipated’ emotional expressions allows us to reliably discriminate real from faked emotions. Having said that, we concur that our work on emotional leakage to date has been conducted exclusively on English-speaking Western populations and hope that cross-cultural research in this area will soon transpire. Although in theory, our work on involuntary emotional leakage of universal emotions should have cross-cultural reliability, empirical attention is needed to address the issue.

10.4 Voluntary and Involuntary Aspects of Emotional Deception

‘This would not be the first time the parole board has been deceived’. This admission came from the National Parole Board of Canada responding to criticisms of its decision to release Robert Moyes despite his horrendous criminal record and psychopathic tendencies (Kari 2006). Upon release, he murdered seven people, and then spoke to the media about how he had ‘cried’ his way out of parole hearings. How could the parole board have been so fooled in accepting his claims of remorse and rehabilitation?

Emotional expressions comprise complex facial muscle activations in both the upper and lower face, some of which are under voluntary control while others are not. Emotional expressions sometimes can be a pertinent source of information about the states and traits of others precisely because they are spontaneous or ‘involuntary’; the onset of expressions occurs rapidly following a strong affective experience and is expressed before the individual is able to consciously process and control them. This overt affective experience can be advantageous, and potentially life-saving, to observers if it helps them to identify a threatening individual in the environment and to act accordingly (Ekman 1997; Hampson et al. 2006; Schmidt and Cohn 2001). However, despite the potential advantage that interpersonal assessment based on the face affords, the evolutionary development of deception – through the manipulation of facial expressions, body language and words – has increased the complexity

of this evaluative process. This might explain the failings of the parole evaluation described at the beginning of this section.

Sometimes, expressed emotion is a salient and true representation of an affective state, but other times, expressed emotion may be a falsified portrayal of an affective state that is not genuinely felt. Unfortunately, it appears that in this co-evolutionary arms race, deceivers have gained the upper hand (ten Brinke and Porter 2013), to the extent that observers typically cannot discriminate sincere from feigned emotions (e.g., Porter and ten Brinke 2008), or truth from lies more generally (Vrij 2008; Vrij, Granhag and Porter 2010). This imbalance has particular ramifications in legal contexts, where relying heavily on the appraisal of emotional communication can result in misguided decisions with high consequences regarding witness credibility, defendant guilt and sentencing (see Bandes, this volume, and Weisman, this volume, for discussion of emotional expression in such high-stakes legal contexts).

Considering the pervasiveness of attending to emotional expressions to make inferences about others which, in turn, informs our behaviour towards them, it may be advantageous for targets to conceal genuine or feign unfelt emotions by manipulating their facial expressions in situations where they wish to mislead observers. Although deceivers may have considerable success in lying to others (see Vrij, Granhag and Porter 2010), they ought to fail more often because their true emotions are 'written all over their face'. In the first detailed examination of what the face reveals during emotional deceit, Porter and ten Brinke (2008) had participants view powerful visual images, responding with a genuine or convincing (but false) expression while being judged by a naïve observer. The nearly seven hundred videotaped expressions were then analyzed (each one-thirtieth of a second frame for more than one hundred thousand frames) by coders blind to expression veracity. Although involuntary expressions were subtle, no one was able to falsify emotions without leakage of their true feelings on at least one occasion. However, the participants were amazingly successful emotional deceivers, and despite the presence of various cues to emotional deceit, the naïve judges were able to discriminate genuine and deceptive expressions at a level only slightly above chance (also see Hess and Kleck 1994). In general, humans are better emotional liars than emotional lie detectors, effectively 'flipping a coin' when asked to evaluate truthfulness (e.g., Bond and DePaulo 2006; Vrij 2008).

10.5 Manifestations of Sincere and Insincere Basic Emotions

The hypothesized basic emotions that we have studied and described earlier serve an adaptive purpose; it is evolutionarily advantageous to experience, communicate and perceive them and their corresponding expressions in others to evaluate others and to facilitate social interactions. As mentioned, early work following Darwin's initial investigation has classified these basic emotions as universal in nature. Again, while acknowledging the controversies around whether these are universal (e.g., Russell 1994, 2003), we have found them to be extremely useful in distinguishing genuine and insincere emotional states in both low-stakes and high-stakes contexts. Here, we turn briefly to what is known about the facial muscles and expressions corresponding to these hypothesized basic emotions.

Happiness is the only positive basic emotion but there are many other related enjoyable emotions, such as joy, awe and pride that result in similar expressions on the face with slight variation (Ekman 2003). The common observable expression that links these pleasurable emotions is the smile on the lower half of the face, an activation of the zygomatic major muscle, pulling the lip corners up and back towards the ears. However, because the smile is common to each of these emotions and social etiquette encourages politeness (e.g., smiling; subsequently giving us practice at effectively activating this muscle and simulating happiness), a smile is not considered to be the sole indicator of genuine happiness. Early work by Duchenne ([1862] 1990) identified a second critical component of the happiness expression through electrical stimulation of various facial muscles. By comparing the facial expression produced by electrical stimulation of the zygomatic major muscle alone and the expression produced when an individual felt genuine happiness, he discovered that a prototypical expression of happiness also includes activation of the orbicularis oculi (the sphincter muscle surrounding the eye), subsequently producing what is now known as a 'Duchenne smile'. This added muscle activation is responsible for creating 'crow's feet' surrounding the eye and together with other minor muscles around the eye creates wrinkling at the outer eye corner, lowering of the eye cover fold and wrinkling of the lower eyelid (Duchenne [1862] 1990; Ekman, Davidson and Friesen 1990). Although a smile can be performed on command, activation of the relevant muscles responsible for a prototypical expression of genuine happiness in the upper face is much

more difficult to spontaneously produce. Thus, a fake smile often will be missing the wrinkling in the corners of the eyes.

The second hypothesized basic facial expression is sadness, an expression that typically is longer lasting than the other basic emotions (e.g., Verduyn and Lavrijsen 2015) and one that is produced following negatively valenced events or experiences. Similar to happiness, the best known or stereotypical indicator of sadness is found on the lower face and is related to muscle activation surrounding the mouth, in this case resulting in a frown (specifically resulting from the activation of the depressor anguli oris; Ekman 2003; Ekman and Friesen 1978). This emotional expression is not complete without the activation of muscles in the upper face as well, specifically the corrugator supercilii between the eyebrows and frontalis muscle in the forehead. Collectively, the corrugator supercilii and frontalis, known as the 'grief muscles', contract and cause the inner eyebrows to move together and upwards. Akin to happiness, people are quite competent at simulating a frown at will, and even at activating the frontalis muscle (moving the forehead upwards only), but typically fail at producing a true sadness expression, which requires activation of the corrugator muscle (pulling the eyebrows together) as well. As such, despite a display of sadness in the lower face, fabricated sadness often resembles a surprised expression in the upper face.

The third hypothesized basic facial expression is surprise, produced in response to sudden, unexpected stimuli. Unlike the other basic emotions, this emotion can be produced in response to either negative or positive stimuli. For example, the same individual may experience surprise in response to encountering a deadly snake or a surprise birthday party. Emotional expressions of surprise are characterized by the raising of the inner and outer eyebrows, widened eyes, open mouth and a slack jaw. In particular, the frontalis muscle in the forehead region becomes activated raising the eyebrows, and the muscles around the mouth (orbicularis oris) and chin become disengaged. Surprise is often experienced for a very brief period of time and acts as a transition emotion that changes quickly after further appraisal of the environment occurs and a different emotion emerges (e.g., to fear in the case of the snake or happiness in the case of the birthday party).

An emotion that observers sometimes confuse with surprise is fear. Fear is a functional emotion because it is triggered by the threat of psychological or physical harm (Grillon and Charney 2011) and induces a fight-or-flight response to inform appropriate action. Some of the

components of the expression of fear are theorized to assist with both fighting and fleeing. For example, fear is expressed on the upper face by the activation of the oblique eyebrow muscles (subsequently producing a frown) and the widening of the eyes which collectively, from a functional explanation, may protect the eyes while allowing them to stay open and track the feared target (Ekman and Friesen 1975; Izard 1986). Fear is displayed in the lower face by drawing back the lips and tightening the sheet of muscle found across the front of the neck (the platysma muscle; Ekman and Friesen 2003). The fear expression also serves a communicative function because it notifies others that there is a threat in the vicinity. Although the two are often confused, each of the fear and surprise expressions has unique features. Fear involves frontalis and corrugator supercilii activation, which pulls the inner eyebrows towards each other, whereas surprise involves the activation of the frontalis muscle only, which results in the vertical elevation of the eyebrows (creating horizontal wrinkles in the forehead). Further, the expression of fear in the lower face involves an active tightening and pulling back of the lips, and the activation of jaw and neck muscles whereas a surprise expression is characterized by a relaxing of the muscles around the mouth allowing the jaw to drop open (i.e., a passive movement).

Of the hypothesized basic emotions, one of the most evolutionarily salient is anger (Amado, Yildirim and Iyilikçi 2011). The various manifestations of anger prepare an individual for action (e.g., fight or flight) if further evaluation of the situation deems it necessary. In the upper face, anger is made apparent by a fierce gaze, which involves a slight widening of the eyes and furrowed brows caused by the activation of the orbicularis oculi around the eyes, the corrugator supercilii and procerus muscles between and below the eyes, in particular. In the lower face, anger involves either the firm pressing together of the lips or the clenching and baring of the teeth, largely requiring activation of the round orbicularis oris sphincter muscle surrounding the mouth (Darwin 1872; Ekman and Friesen 2003).

The sixth hypothesized basic facial expression is disgust. Disgust is a visceral reaction to something that is offensive to the senses and can be triggered by noxious stimuli, such as faeces, bacteria and parasites, or disturbing images. This emotion can arise if the observer is directly exposed to the offensive trigger or witnesses others being exposed to it. Expressions of disgust in the upper face involve lowering of the eyebrows and wrinkling of the nose, and drawing upwards of the upper

lip (typically exposing the teeth; Ekman and Friesen 2003) in the lower face – the hallmark of this expression. Similar to fear, disgust is an adaptive experience that draws attention to potentially threatening, toxic stimuli in the environment, and ensures that they are avoided (e.g., Oaten, Stevenson and Case 2009). As such, this expression is easily simulated but is very difficult to mask.

The seventh, final and still debated basic expression is contempt, or the experience of animosity or condescension towards others (Matsumoto 1992). It is felt most often in situations in which one individual feels superior to and looks down upon another. Contempt often manifests unilaterally on the face, that is, it is expressed on only one half of the face (i.e., upwards pull of only one lip corner). While the upper half of the face is not significantly involved in the expression of contempt, the lower half of the face reveals contempt when the individual's lips are pressed tightly together and one side of the upper lip is raised, requiring involvement of the orbicularis oris (Ekman and Friesen 2003). There is less support for contempt as a universal expression because the interpretation of this emotion has been shown to be context dependent (e.g., Russell 1991); that is, people have difficulty labelling the expression without an environmental or situational prompt (Rosenberg and Ekman 1995).

Again, despite support for the existence of these basic emotions, some scholars suggest a larger influence of culture exists than has been appreciated to date (Jack et al. 2009; Jack, Caldara and Schyns 2012; see also Jack, this volume; Russell, this volume). More specifically, some posit that emotional expression can be influenced by culture-specific interpretations that influence the processing of facial expressions and other non-verbal behaviours. In sum, the results of this research remain mixed, with support for both cultural specificity and the universality of expressions.

10.5.1 Voluntary Aspects of Insincere Emotional Expressions

Over time, humans have evolved to become increasingly proficient at controlling some aspects of facial expressions to facilitate and improve social interactions. There are many situations in which displaying genuine emotions on the face could be detrimental. For example, someone might conceal his/her true emotions for altruistic reasons; a guest may hide his/her dislike of their host's meal with a smile, or an individual undergoing severe medical treatments may conceal his/her agony and

put on a happy face when family members visit to reduce their concern (see Gross and Thompson 2007 for discussion). Although this may be selfishly motivated at times, it also allows for people to thrive in the complex social environment in which we live. There is also evidence for the benefits of adopting initially unfelt emotional expressions; people who smile begin to feel happier after doing so (Kraft and Pressman 2012). However, humans also manipulate their emotional expressions for malicious and/or antisocial purposes, such as attempting to deceive others.

Regardless of the motivation for doing so, there are three ways that emotional facial expressions are intentionally manipulated (Ekman and Friesen 2003): simulating, masking and neutralizing. Simulating involves adopting an emotional expression in the absence of any real emotion; masking involves replacing a felt emotion with a different, unfelt emotional expression; and neutralizing involves concealing a felt emotion with a neutral facial expression. For example, someone trying to hide their genuine feelings of happiness in response to another's misfortune can choose to either neutralize their expression or mask it with a more socially appropriate display of sympathy. Alternatively, if an individual feels no emotion in response to the other's misfortune but is aware of the socially acceptable response, they may simulate a display of empathy.

Despite the value of being able to manipulate and falsify emotional expressions, this often is difficult because some facial muscles are more easily controlled than others. The muscles involved in the facial expressions of the basic emotions that are under greater voluntary control include the frontalis (moving eyebrows up and down), depressor anguli oris (pulling mouth corners downwards) and zygomatic major (pulling mouth corners upwards). These muscles are easily manipulated and can be used to portray one component of an emotional expression (e.g., moving the frontalis upwards contributes to an expression of surprise in the upper half of the face and contracting the zygomatic major muscle results in a smile in the lower half of the face). Conversely, other facial muscles such as the corrugator supercilii (pulling eyebrows together) and orbicularis oculi (sphincter muscle around the eye to induce squinting) are more difficult to consciously control (Ekman 2003; ten Brinke et al. 2012).

The degree to which a muscle is under voluntary control appears to be the result of neural substrates; that is, facial muscles that can be intentionally contracted may have greater representation in the

motor cortex of the brain (see Ekman 2003 for a discussion). Further, it appears that the emotional displays most commonly required to meet social obligations, and thus most commonly displayed, are more easily manipulated (such as the socially expected polite smile). Because each emotional expression comprises a multitude of muscles working together, the activation of isolated muscles does not necessarily result in the successful simulation of a genuine emotional expression. This means that an individual simulating happiness may produce a believable smile but will likely be missing the accompanying muscle activations around the eye, such as the orbicularis oculi, required to produce a prototypical Duchenne smile (Frank, Ekman and Friesen 1993).

10.5.2 *Involuntary Aspects of Insincere Emotional Expressions*

Despite the various strategies for fabricating insincere emotions or concealing genuine ones, a long-standing theory suggests that true emotions are revealed on the face involuntarily and despite a target's best efforts. The first to consider this was Darwin (1872) who posited that certain facial muscles cannot be (1) intentionally activated in the absence of genuine emotion (as per the discussion earlier), or (2) suppressed in the presence of a genuine emotion, especially those that are potent. Collectively, these two propositions form the *inhibition hypothesis* (Darwin 1872; Ekman 2003). Although this proposal has major theoretical and applied implications, it had received little empirical scrutiny until recently (e.g., Hurley and Frank 2011; Porter and ten Brinke 2008; Porter, ten Brinke and Wallace 2012). Since being subjected to critical evaluation, research has consistently found support for Darwin's (1872) original assertions. For example, as described earlier, Porter and ten Brinke (2008) presented participants with powerful emotional images to induce various basic discrete emotions and then sometimes instructed participants to adopt expressions inconsistent with the induced emotion. The results indicated that not a single participant was able to wholly mask his/her emotions and instead displayed emotional leakage, particularly under high emotional intensity (see also Porter, ten Brinke and Wallace 2012).

The ability to simulate and conceal genuine emotional expressions also has been tested in a high-stakes, real-world setting. Specifically, in the largest-scale study of high-stakes emotional deception to date, we (ten Brinke and Porter 2012; ten Brinke, Porter and Baker 2012) found support for both prongs of the inhibition hypothesis. Coding the

emotional expressions of individuals televised while pleading for the safe return of a missing family frame by frame, we found evidence for the inability to conceal felt emotions, demonstrated by expressions of fear, disgust, smirks and anger on the face of deceptive (those responsible for the missing individuals' death) but not the honest pleaders. Similarly, there was evidence for the inability to fully simulate unfelt emotions in the failed attempts to express sadness (i.e., lack of corrugator contraction) on the faces of deceptive pleaders only. The *inhibition hypothesis* (Darwin 1872; Ekman 2003) continues to be dominant in the field of emotional expression and guides our understanding of the manner in which, and reason why, emotion is 'leaked' on the face.

Despite discovering support for the inhibition hypothesis, our understanding of the manner in which emotional expressions manifest involuntarily has developed significantly since Darwin's time and has been a popular topic within scientific and media arenas (e.g., Adelson 2004; Duenwald 2005; Ekman 2006; Henig 2006; Schubert 2006). Ekman (1992a, 2006) extended Darwin's hypothesis to suggest that when an emotion is concealed, emotional leakage of the felt emotion manifests in the form of a 'micro-expression', an observable, but fleeting manifestation ($1/25$ – $1/5$ of a second) of an emotional expression that appears across the entire face. At the time of this proposal, there was no empirical support for this phenomenon, and the nature of micro-expressions has since been questioned. For example, research has found that these fleeting expressions do not manifest across the entire face but, rather, are expressed in either the upper or lower face (e.g., Ekman 1992b; Frank and Svetieva 2015; Porter and ten Brinke 2008). Further, more recent research suggests that emotional leakage (i.e., micro-expressions) tend to last much longer than initially hypothesized (Ekman 1992b), as long as one full second (e.g., Porter and ten Brinke 2008; Porter et al. 2012; Yan et al. 2013). Yan et al. (2013) have proposed that a micro-expression should be defined as any emotional expression lasting less than 500 ms. Further, previous studies have found variation in the duration, onset and offset times between genuine and false expressions of happiness and disgust (e.g., Frank, Ekman and Friesen 1993; Hess and Kleck 1990). Despite the absence of full-face, tell-tale micro-expressions, there certainly is growing evidence for the presence of involuntary emotional leakage on the face that provides insight into a target's true emotions, or at least their emotional insincerity.

Although it may be the case that no one is immune to emotional leakage on the face, it appears that some individuals are particularly adept

at facial control. Porter et al. (2012) hypothesized that two individual differences that may relate to 'expertise' in the communication of facial expressions (i.e., the ability to adopt convincing emotional displays and avoid the problem of emotional 'leakage') were emotional intelligence (EI) and psychopathy. Those with high EI are skilled at identifying and managing their own emotions and identifying the emotional expressions of others (Petrides, Perez-Gonzalez and Furnham 2007). Conversely, those scoring high in psychopathy possess an emotional deficit, a dearth of emotional experience (e.g., Herpertz et al. 2001), but the ability to mimic emotional expressions to manipulate others (Book et al. 2015). It is theorized that those with high EI would demonstrate emotional mastery because they are in tune with their emotions and knowledgeable about emotional expression whereas psychopathy was theorized to be related to emotional expression control because psychopaths have practice in simulating emotions and would not have to suppress competing emotions, as a result of having less of an emotional reaction to the stimuli. To test this theory, participants were presented with images to induce discrete emotions (happiness, sadness, disgust and fear) and were instructed to respond genuinely or with another, discrete emotional expression. Following frame-by-frame analysis used to code for the contraction of facial muscles related to emotional expressions and apparent emotional leakage, both individual difference groups showed prowess in the manipulation of emotional expressions in line with the hypothesis. Specifically, those scoring high in EI – and in a subscale measuring the ability to perceive and express emotion more specifically – feigned emotions more convincingly than others (i.e., were better able to display consistent emotional expressions in both genuine and deceptive emotional displays). Further, psychopathic traits were related to shorter durations of unintended emotional 'leakage' (a finding recently replicated in criminal offenders by Book et al. 2015). Other combinations of personality traits, such as high self-monitoring and low sincerity, also may be beneficial to emotional control (Grieve 2011).

10.6 Communicating False Emotion in High-Stakes Contexts

The communication of emotion on the face, particularly as it relates to credibility assessment, is relevant at all stages of the legal system. From the interrogation of suspects and confessions (e.g., Hargrave 2008), courtroom testimony (e.g., Kaufmann et al. 2003), through to

sentencing and parole decisions, the face is relied upon as an indicator of a target's degree of sincerity. This is made evident by statements by Canadian judges such as 'judges and jurors must view a witness to adequately evaluate body language, facial expressions, and other indicators of credibility' (*R. v. B. (K.G.)* 1993). Further, this view has been enforced by the decision to require women to remove their niqabs (a religious garment that obscures a woman's face) before testifying in a courtroom because they 'impair a judge's ability to assess credibility' and impede examination because the lawyers cannot re-direct their line of questioning based on demeanour (e.g., *The Star* 2009).

Owing to society's expectations that emotional displays should be congruent with the circumstances, offenders may be motivated to display expressions of sadness or distress while discussing their crimes, sometimes insincerely. This is in line with Affect Control Theory (ACT; Robinson and Smith-Lovin 1999), according to which people attempt to influence others' impressions of them by exhibiting expressions congruent with society's expected schemas for particular situations. Indeed, research has found that people who exhibit the expected and approved emotions in relation to their actions receive positive interpersonal evaluations. They are rated highest on being nice, good, sincere, genuine and likeable (e.g., Robinson and Smith-Lovin 1999) and are much more likely to be forgiven for their transgressions (Davis and Gold 2011; Gold and Weiner 2000). As such, in addition to influencing interpersonal perception at the individual-level, emotional expression in the legal system plays a very important communicative role (see Sorial, this volume, for a discussion of one such role). For example, remorse is a particularly important moral emotion that is expected in legal contexts and is associated with favourable outcomes in the legal system, such as reduced sentences and early release (see Weisman, this volume). Indeed, the relevance of remorse during sentencing, treatment prognosis and rehabilitation has been explicitly noted in a number of legal cases (e.g., *R. v. Struve* 2007; see also Slovenko 2006) and empirical studies (MacLin et al. 2009; Pipes and Alessi 1999; Taylor and Kleinke 1992; Wiener and Reinehart 1986). Because of the weight given to perceptions of remorse in legal contexts, perpetrators lacking remorse can be highly motivated to feign it to garner undeserved benefits of the emotional display. To account for the communicative function of the expression of remorse, two of the present authors (as part of the first author's doctoral dissertation) propose the Dual-Process Functional Theory (DPFT) of remorse founded on related functional frameworks of emotion (e.g.,

Ekman 1992a; Horberg et al. 2011; Keltner and Gross 1999), emphasizing the need to examine emotions from a social functional perspective. In the legal system, the explicit communication of emotion to others within the social group allows for observers to discern that the transgressor understands the violation of social standards and rebuilds the previously held trust within the community, avoiding ostracism and enhancing the survival of the transgressor and insularity/strength of the group (see Helm, this volume, for a discussion of the social role of the 'reactive attitudes' and Weisman, this volume, for a discussion of the specific role of remorse).

Despite the motivation to communicate remorse, when individuals attempt to display remorse without accompanying feelings of remorse, differences in emotional expression can be detected. For example, ten Brinke et al. (2012) found that true and feigned remorse displays have distinctive features that are salient to the informed observer, such that false remorse is associated with more neutral expressions and a mix of brief expressions of surprise and happiness. Conversely, genuine remorse was associated with primarily sad but also neutral expressions, and rarely did other emotional expressions emerge.

Although remorse serves as a particularly salient example of emotional expression in the legal system, displays of basic emotions also serve an important communicative role, and an attuned observer can observe differences between genuine and falsified emotional displays. Cues to one's actual emotions may be particularly salient in the legal system because of increased cognitive load (Vrij 2008) allowing observers the opportunity to scrutinize the face of targets to inform their decision making. More specifically, cues from displayed emotions can be incorporated into information gleaned from other communication channels, such as non-verbal displays (e.g., use of illustrators to distract from speech content) and verbal aspects (e.g., psychological distancing through pronoun use), and can be used to inform credibility assessments in a holistic manner (considered to be the best approach for detecting deception and associated with higher overall accuracy rates; e.g., Ekman and O'Sullivan 1991; Vrij, Akehurst, Soukara and Bull 2004; Vrij, Evans, Akehurst and Mann 2004). However, information gathered from an evaluation of one's emotional expression on the face alone can still be valuable.

Relying on emotional expression to inform decision making and subsequent interactions is particularly important because we know that individuals often are less than forthcoming, especially in cases

of unfavourable emotional experience. For example, an emotional experience gaining popularity in the scientific literature is *schadenfreude*. *Schadenfreude* is pleasure derived from another person's misfortune (but differs from sadism in that the individual did not play a direct role in that person's misfortune). This emotional experience is more likely to be elicited under certain circumstances, such as when the misfortune happens to someone who is envied (e.g., Smith et al. 1996), disliked (e.g., Hareli and Weiner 2002) or deserving (e.g., Van Dijk et al. 2005). Although typically displayed in social situations (such as when viewing a comical 'America's Funniest Home Videos' episode), displays of *schadenfreude* would be considered inappropriate in a legal context. For example, it is concerning when *schadenfreude* is experienced by legal decision makers such as prosecutors who derive pleasure from the misfortunes suffered by the defendant or joy derived from the harsh sentencing of a defendant. However, the emotional experience and expression of *schadenfreude* highlights the value of attending to the face. Porter et al. (2014) found that although individuals fail to self-report experiences of *schadenfreude*, a more objective evaluation of these participants' emotional cues suggest that they experience it (with brief flashes of happiness being apparent on the face). Subsequently, because of the emphasis placed on appropriate emotional displays in line with certain situations, emotional leakage provides objective information about others' inner states.

Although cues to others' emotional states do not serve as a Pinocchio's nose (i.e., a smoking gun in the deception realm), information about emotions gleaned from the face can be used to guide subsequent interactions. For example, in social interactions, observers can use emotional information to read between the lines and identify affect that the target may not be forthcoming about, whether in a conversation with a significant other or a business negotiation. In a legal context, one's revealed emotions can guide further questioning in an investigative context. However, we must concede that observers typically are quite poor at evaluating the sincerity of others' emotional displays. For example, in the previous studies finding support for the inhibition hypothesis (Porter and ten Brinke 2008; Porter et al. 2012), and thus, the presence of observable cues to emotional deception, naïve observers were unable to correctly differentiate sincere from insincere emotional displays. Despite the fact that observers are typically poor at reading the face, accuracy can be improved with training focused on increasing

one's understanding and ability to detect genuine, and falsified, emotional expressions (e.g., Shaw, Porter and ten Brinke 2013).

10.7 Conclusion and Future Directions

In summary, theory and research collectively emphasize the importance of the human face in revealing covert emotional information. Following from the inhibition hypothesis (Darwin 1872; Ekman 2003), our research supports the contention that people cannot completely inhibit powerfully felt emotions and are likely to fail when attempting to falsify these expressions in the absence of emotion. In both controlled research settings and real-life contexts, emotional leakage occurs that can be identified by the informed observer. Indeed, the secrets of the human face finally are being revealed empirically. Future research needs to address whether the types of involuntary facial expression 'leakage' found with English-speaking Western targets has cross-cultural validity. Also, more research is needed to examine whether observers can be trained to identify emotional leakage (e.g., Shaw, Porter and ten Brinke 2013), and whether automated software can be created to effectively and reliably evaluate the human face for emotional leakage.

Acknowledgements

This chapter and many of the authors' research discussed in it were funded by grants from the Social Sciences and Humanities Research Council of Canada (SSHRC).

References

- Abdel-Rahman, R. 2011. 'Facing Good and Evil: Early Brain Signatures of Affective Biographical Knowledge in Face Recognition', *Emotion* 11: 1397–1405.
- Adams, R. B., Ambady, N., Macrae, C. N. and Kleck, R. E. 2006. 'Emotional Expressions Forecast Approach-Avoidance Behavior', *Motivation and Emotion* 30: 177–186.
- Adelson, R. 2004. 'Detecting Deception', *APA Monitor on Psychology* 35: 70–73.
- Amado, S., Yildirim, T. and Iyilikçi, O. 2011. 'Observer and Target Sex Differences in the Change Detection of Facial Expressions: A Change Blindness Study', *Cognition, Brain, Behavior* 15: 295–316.
- Amihai, I., Deouell, L. Y. and Bentin, S. 2011. 'Neural Adaptation Is Related to Face Repetition Irrespective of Identity: A Reappraisal of the N170 Effect', *Experimental Brain Research* 209: 193–204.

- Arsalidou, M., Morris, D. and Taylor, M. J. 2011. 'Converging Evidence for the Advantage of Dynamic Facial Expressions', *Brain Topography* 24: 149–163.
- Barrett, L. F. 1998. 'Discrete Emotions or Dimensions? The Role of Valence Focus and Arousal Focus', *Cognition and Emotion* 12: 579–599.
- Barrett, F. L. and Russell, J. A. 1998. 'Independence and Bipolarity in the Structure of Current Affect', *Journal of Personality and Social Psychology* 74: 967–984.
- Barrett, L. F., Lindquist, K. A. and Gendron, M. 2007. 'Language as Context for the Perception of Emotion', *Trends in Cognitive Sciences* 11: 327–332.
- Bond, C. F. and DePaulo, B. M. 2006. 'Accuracy of Deception Judgments', *Personality and Social Psychology Review* 10: 214–234.
- Bond, C. F., Berry, D. S. and Omar, A. 1994. 'The Kernel of Truth in Judgments of Deceptiveness', *Basic and Applied Social Psychology* 15: 523–534.
- Book, A., Methot, T., Gauthier, N., Hosker-Field, A., Forth, A., Quinsey, V. and Molnar, D. 2015. 'The Mask of Sanity Revisited: Psychopathic Traits and Affective Mimicry', *Evolutionary Psychological Science* 1: 91–102.
- Borkenau, P., and Liebler, A. 1992. 'Trait Inferences: Sources of Validity at Zero Acquaintance', *Journal of Personality and Social Psychology* 62: 645–657.
- Darwin, C. 1872. *The Expression of the Emotions in Man and Animals*. New York: Oxford University Press.
- Davis, J. R. and Gold, G. J. 2011. 'An Examination of Emotional Empathy, Attributions of Stability, and the Link between Perceived Remorse and Forgiveness', *Personality and Individual Differences* 50: 392–397.
- Duchenne, G. B. [1862] 1990. *The Mechanism of Human Facial Expression*. New York: Cambridge University Press.
- Duenwald, M. 2005. 'The Physiology of Facial Expressions'. <http://discovermagazine.com/2005/jan/physiology-of-facial-expressions>
- Ekman, P. 1980. *Face of Man: Universal Expression in a New Guinea Village*. New York: Garland.
- 1992a. 'An Argument for Basic Emotions', *Cognition and Emotion* 6: 169–200.
- 1992b. *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*. New York: Norton.
1993. 'Facial Expression and Emotion', *American Psychologist* 48: 384–392.
1997. 'Expression or Communication about Emotion', in Segal, Weisfeld and Weisfeld (eds.) *Genetic, Ethological and Evolutionary Perspectives on Human Development: Essays in Honor of Dr. Daniel G. Freedman*. Washington, DC: American Psychological Association.
2003. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. New York: St. Martin's Griffin.
2006. 'How to Spot a Terrorist on the Fly'. www.washingtonpost.com, retrieved 29 October 2006.
- Ekman, P. and Cordaro, D. 2011. 'What Is Meant by Calling Emotions Basic', *Emotion Review* 3: 364–370.
- Ekman, P. and Friesen, W. V. 1971. 'Constants across Cultures in the Face and Emotion', *Journal of Personality and Social Psychology* 17: 124–129.
1978. *The Facial Action Coding System*. Palo Alto, CA: Consulting Psychologists Press.

2003. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Cues*. Englewood Cliffs, NJ: Prentice-Hall.
- Ekman, P. and Oster, H. 1979. 'Facial Expressions of Emotion', *Annual Review of Psychology* 30: 527–554.
- Ekman, P. and O'Sullivan, M. 1991. 'Who Can Catch a Liar?' *American Psychologist* 46: 913–920.
- Ekman, P., Davidson, R. J. and Friesen, W. V. 1990. 'The Duchenne Smile: Emotional Expression and Brain Physiology: II', *Journal of Personality and Social Psychology* 58: 342–353.
- Ekman, P., Friesen, W. V. and Ellsworth, P. 1972. *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. Oxford: Pergamon Press.
- Ekman, P., Sorenson, E. R. and Friesen, W. V. 1969. 'Pan-Cultural Elements in Facial Displays of Emotion', *Science* 164: 86–88.
- Fox, C. J., Hanif, H. M., Iaria, G., Duchaine, B. C. and Barton, J. J. 2011. 'Perceptual and Anatomic Patterns of Selective Deficits in Facial Identity and Expression Processing', *Neuropsychologia* 49: 3188–3200.
- Frank, M. G. and Svetieva, E. 2015. 'Microexpressions and Deception', in Mandal and Awasthi (eds.) *Understanding Facial Expressions in Communication*. New York: Springer, pp. 227–242.
- Frank, M. G., Ekman, P. and Friesen, W. V. 1993. 'Behavioral Markers and Recognizability of the Smile of Enjoyment', *Journal of Personality and Social Psychology* 64: 83–93.
- Gold, G. J. and Weiner, B. 2000. 'Remorse, Confession, Group Identity, and Expectancies about Repeating a Transgression', *Basic and Applied Social Psychology* 22: 291–300.
- Goldfield, E. C. 1995. *Emergent Forms: Origins and Early Development of Human Action and Perception*. New York: Oxford University Press.
- Grieve, R. 2011. 'Mirror Mirror: The Role of Self-Monitoring and Sincerity in Emotional Manipulation', *Personality and Individual Differences* 51: 981–985.
- Grillon, C. and Charney, D. R. 2011. 'In the Face of Fear: Anxiety Sensitizes Defensive Responses to Fearful Faces', *Psychophysiology* 48: 1745–1752.
- Gross, J. J. and Thompson, R. A. 2007. 'Emotion Regulation: Conceptual Foundations', in Gross (ed.) *Handbook of Emotion Regulation*. New York: Guilford Press, pp. 3–24.
- Hamann, S. 2012. 'What Can Neuroimaging Meta-Analyses Really Tell Us about the Nature of Emotion?', *Behavioral and Brain Sciences* 35: 150–152.
- Hampson, E., van Anders, S. M. and Mullin, L. I. 2006. 'A Female Advantage in the Recognition of Emotional Facial Expressions: Test of an Evolutionary Hypothesis', *Evolution and Human Behavior* 27: 401–416.
- Hareli, S. and Weiner, B. 2002. 'Dislike and Envy as Antecedents of Pleasure at Another's Misfortune', *Motivation and Emotion* 26: 257–277.
- Hargrave, J. 2008. 'Do You Speak Body Language?: Mastering the Art of Nonverbal Communication Key in Interrogations', *The Forensic Examiner* 17: 17–22.
- Henig, R. M. 2006. 'Looking for the Lie'. www.nytimes.com, retrieved 19 September 2007.

- Herpertz, S. C., Werth, U., Lukas, G., Qunaibi, M., Schuerkens, A., Kunert, H., Freese, R., Flesch, M., Mueller-Isberner, R., Osterheider, M. and Sass, H. 2001. 'Emotion in Criminal Offenders with Psychopathy and Borderline Personality Disorder', *Archives of General Psychiatry* 58: 737–745.
- Hess, U. and Kleck, R. E. 1990. 'Differentiating Emotion Elicited and Deliberate Emotional Facial Expressions', *European Journal of Social Psychology* 20: 369–385.
1994. 'The Cues Decoders Use in Attempting to Differentiate Emotion-Elicited and Posed Facial Expressions', *European Journal of Social Psychology* 24: 367–381.
- Horberg, E. J., Oveis, C. and Keltner, D. 2011. 'Emotions as Moral Amplifiers: An Appraisal Tendency Approach to the Influences of Distinct Emotions upon Moral Judgment', *Emotion Review* 3: 237–244.
- Hurley, C. M. and Frank, M. G. 2011. 'Executing Facial Control during Deception Situations', *Journal of Nonverbal Behavior* 35: 119–131.
- Izard, C. E. 1986. *The Emotions*. Cambridge: Cambridge University Press.
1994. 'Innate and Universal Facial Expressions: Evidence from Developmental and Cross-Cultural Research', *Psychological Bulletin* 115: 288–299.
- Jack, R. E., Caldara, R. and Schyns, P. G. 2012. 'Internal Representations Reveal Cultural Diversity in Expectations of Facial Expressions of Emotion', *Journal of Experimental Psychology: General* 141: 19–25.
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G. and Caldara, R. 2009. 'Cultural Confusions Show That Facial Expressions Are Not Universal', *Current Biology* 19: 1543–1548.
- Jack, R. E., Garrod, O. G., Yu, H., Caldara, R. and Schyns, P. G. 2012. 'Facial Expressions of Emotion Are Not Culturally Universal', *Proceedings of the National Academy of Sciences of the United States of America* 109: 7241–7244.
- Kanwisher, N., McDermott, J. and Chun, M. M. 1997. 'The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception', *The Journal of Neuroscience* 17: 4302–4311.
- Kari, S. 2006. 'Probe Finds No Fault in Killer's Parole: Career Criminal Took Part in Seven Murders While on "Conditional Release"'. Web site. URL no longer available.
- Kaufmann, G., Drevland, G. C., Wessel, E., Overskeid, G. and Magnussen, S. 2003. 'The Importance of Being Earnest: Displayed Emotions and Witness Credibility', *Applied Cognitive Psychology* 17: 21–34.
- Kayyal, M. H. and Russell, J. A. 2013. 'Americans and Palestinians Judge Spontaneous Facial Expressions of Emotion', *Emotion* 13: 891–904.
- Keltner, D. and Gross, J. J. 1999. 'Functional Accounts of Emotions', *Cognition and Emotion* 13: 467–480.
- Kraft, T. L. and Pressman, S. D. 2012. 'Grin and Bear It: The Influence of Manipulated Facial Expression on the Stress Response', *Psychological Science* 23: 1372–1378.
- Kuppens, P., Tuerlinckx, F., Russell, J. A. and Barrett, L. F. 2013. 'The Relation between Valence and Arousal in Subjective Experience', *Psychological Bulletin* 139: 917–940.

- MacLin, M. K., Downs, C., MacLin, O. H. and Caspers, H. M. 2009. 'The Effect of Defendant Facial Expression on Mock Juror Decision-Making: The Power of Remorse', *North American Journal of Psychology* 11: 323–332.
- Matsumoto, D. 1992. 'American-Japanese Cultural Differences in the Recognition of Universal Facial Expressions', *Journal of Cross-Cultural Psychology* 23: 72–84.
- Nelson, N. L. and Russell, J. A. 2013. 'Universality Revisited', *Emotion Review* 5: 8–15.
- Oaten, M., Stevenson, R. J. and Case, T. I. 2009. 'Disgust as a Disease-Avoidance Mechanism', *Psychological Bulletin* 135: 303–321.
- Oosterhof, N. N. and Todorov, A. 2008. 'The Functional Basis of Face Evaluation', *Proceedings of the National Academy of Sciences of the United States of America* 105: 11087–11092.
- Ortega, J. E., Iglesias, J., Fernández, J. M. and Corraliza, J. A. 1983. 'La Expresión Facial en los Ciegos Congénitos (Facial Expression among the Congenitally Blind)', *Infancia y Aprendizaje/Journal for the Study of Education and Development* 21: 83–96.
- Petrides, K. V., Perez-Gonzalez, J. C. and Furnham, A. 2007. 'On the Criterion and Incremental Validity of Trait Emotional Intelligence', *Cognition and Emotion* 21: 26–55.
- Pipes, R. B. and Alessi, M. 1999. 'Remorse and a Previously Punished Offense in Assignment of Punishment and Estimated Likelihood of a Repeated Offense', *Psychological Reports* 85: 246–248.
- Porter, S. and ten Brinke, L. 2008. 'Reading between the Lies: Identifying Concealed and Falsified Emotions in Universal Facial Expressions', *Psychological Science* 19: 508–514.
2009. 'Dangerous Decisions: A Theoretical Framework for Understanding How Judges Assess Credibility in the Courtroom', *Legal and Criminological Psychology* 14: 119–134.
- Porter, S., ten Brinke, L. and Wallace, B. 2012. 'Secrets and Lies: Involuntary Leakage in Deceptive Facial Expressions as a Function of Emotional Intensity', *Journal of Nonverbal Behavior* 36: 23–37.
- Porter, S., ten Brinke, L. and Wilson, K. 2009. 'Crime Profiles and Conditional Release Performance of Psychopathic Sexual Offenders', *Legal and Criminological Psychology* 14: 554–562.
- Porter, S., Bhanwer, A., Woodworth, M. and Black, P. J. 2014. 'Soldiers of Misfortune: An Examination of the Dark Triad and the Experience of Schadenfreude', *Personality and Individual Differences* 67: 64–68.
- Porter, S., ten Brinke, L., Baker, A. and Wallace, B. 2012. 'Would I Lie to You? "Leakage" in Deceptive Facial Expressions to Psychopathy and Emotional Intelligence', *Personality and Individual Differences* 51: 133–137.
- Porter, S., England, L., Juodis, M., ten Brinke, L. and Wilson, K. 2008. 'Is the Face the Window to the Soul?: Investigation of the Accuracy of Intuitive Judgments of the Trustworthiness of Human Faces', *Canadian Journal of Behavioural Science* 40: 171–177.
- R. v. B. (K. G.) 1993. 1 S.C.R. 740.

- R. v. Struve. 2007. BCSC 1316.
- Robinson, D. T. and Smith-Lovin, L. 1999. 'Emotion Display as a Strategy for Identity Negotiation', *Motivation and Emotion* 23: 73–104.
- Rosenberg, E. L. and Ekman, P. 1995. 'Conceptual and Methodological Issues in the Judgment of Facial Expressions of Emotion', *Motivation and Emotion* 19: 111–138.
- Russell, J. A. 1991. 'The Contempt Expression and the Relativity Thesis', *Motivation and Emotion* 15: 149–168.
1994. 'Is There Universal Recognition of Emotion from Facial Expressions? A Review of the Cross-Cultural Studies', *Psychological Bulletin* 115: 102–141.
2003. 'Core Affect and the Psychological Construction of Emotion', *Psychological Review* 110: 145–172.
- Sauter, D., Eisner, F., Ekman, P. and Scott, S. K. 2009. 'Universal Vocal Signals of Emotion', in Van Rijn (ed.) *Proceedings of the 31st Annual Meeting of the Cognitive Science Society*. Amsterdam: Cognitive Science Society, pp. 2251–2255.
- Scherer K. R. and Ekman, P. 1984. *Approaches to Emotion*. Hillsdale, NJ: Erlbaum.
- Schmidt, K. L. and Cohn, J. F. 2001. 'Human Facial Expressions as Adaptations: Evolutionary Questions in Facial Expression Research', *American Journal of Physical Anthropology* 116: 3–24.
- Schubert, S. 2006. 'A Look Tells All', *Scientific American Mind* 17: 26–31.
- Shaw, J., Porter, S. and ten Brinke, L. 2013. 'Catching Liars: Training Mental Health and Legal Professionals to Detect High-Stakes Lies', *The Journal of Forensic Psychiatry & Psychology* 24: 145–159.
- Sievers, B., Polansky, L., Casey, M. and Wheatley, T. 2013. 'Music and Movement Share a Dynamic Structure That Supports Universal Expressions of Emotion', *Proceedings of the National Academy of Sciences of the United States of America* 110: 70–75.
- Slovenko, R. 2006. 'Remorse', *Journal of Psychiatry & Law* 34: 397–432.
- Smith, R., Turner, T., Garonzik, R., Leach, C., Urch-Druskat, V. and Weston, C. 1996. 'Envy and Schadenfreude', *Personality and Social Psychology Bulletin* 22: 158–168.
- Stirrat, M. and Perrett, D. I., 2010. 'Valid Facial Cues to Cooperation and Trust: Male Facial Width and Trustworthiness', *Psychological Science* 21: 349–354.
- Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W. and Anderson, A. K. 2008. 'Expressing Fear Enhances Sensory Acquisition', *Nature Neuroscience* 11: 843–850.
- Taylor, C. and Kleinke, C. L. 1992. 'Effects of Severity of Accident, History of Drunk Driving, Intent, and Remorse on Judgments of a Drunk Driver', *Journal of Applied Social Psychology* 22: 1641–1655.
- ten Brinke, L. and Porter, S. 2012. 'Cry Me a River: Identifying the Behavioral Consequences of Extremely High-Stakes Interpersonal Deception', *Law and Human Behavior* 36: 469–477.
2013. 'Discovering Deceit: Applying Laboratory and Field Research in the Search for Truthful and Deceptive Behavior', in Cooper, Griesel and Ternes (eds.) *Applied Issues in Investigative Interviewing, Eyewitness Memory, and Credibility Assessment*. New York: Springer, pp. 221–237.

- ten Brinke, L., Porter, S. and Baker, A. 2012. 'Darwin the Detective: Observable Facial Muscle Contractions Reveal Emotional High-Stakes Lies', *Evolution and Human Behavior* 33: 411–416.
- ten Brinke, L., MacDonald, S., Porter, S. and O'Connor, B. 2012. 'Crocodile Tears: Facial, Verbal, and Body Language Behaviors Associated with Genuine and Fabricated Remorse', *Law and Human Behavior* 36: 51–59.
- The Star 2009. 'Order to Take off Niqab Pits Law against Religion'. www.thestar.com/news/gta/2009/02/02/order_to_take_off_niqab_pits_law_against_religion.html, retrieved 2 February 2009.
- Valenza, E., Simion, F., Cassia, V. M. and Umiltà, C. 1996. 'Face Preference at Birth', *Journal of Experimental Psychology: Human Perception and Performance* 22: 892–903.
- Van Dijk, W. W., Ouwerkerk, J. W., Goslinga, S. and Nieweg, M. 2005. 'Deservingness and Schadenfreude', *Cognition and Emotion* 19: 933–939.
- Verduyn, P. and Lavrijsen, S. 2015. 'Which Emotions Last Longest and Why: The Role of Event Important and Rumination', *Motivation and Emotion* 39: 119–127.
- Vrij, A. 2008. *Detecting Lies and Deceit: Pitfalls and Opportunities*. Chichester, UK: John Wiley & Sons.
- Vrij, A., Granhag, P. A. and Porter, S. 2010. 'Pitfalls and Opportunities in Nonverbal and Verbal Lie Detection', *Psychological Science in the Public Interest* 11: 89–121.
- Vrij, A., Akehurst, L., Soukara, S. and Bull, R. 2004. 'Let Me Inform You How to Tell a Convincing Story: CBCA and Reality Monitoring Scores as a Function of Age, Coaching, and Deception', *Canadian Journal of Behavioural Science* 36: 113–126.
- Vrij, A., Evans, H., Akehurst, L. and Mann, S. 2004. 'Rapid Judgments in Assessing Verbal and Nonverbal Cues: Their Potential for Deception Researchers and Lie Detection', *Applied Cognitive Psychology* 18: 283–296.
- Waller, B. M. and Micheletta, J. 2013. 'Facial Expression in Nonhuman Animals', *Emotion Review* 5: 54–59.
- Wiener, R. L. and Rinehart, N. 1986. 'Psychological Causality in the Attribution of Responsibility for Rape', *Sex Roles* 14: 369–382.
- Willis, J. and Todorov, A. 2006. 'First Impressions: Making Up Your Mind after a 100-ms Exposure to a Face', *Psychological Science* 17: 592–598.
- Yan, W., Wu, Q., Liang, J., Chen, Y. and Fu, X. 2013. 'How Fast Are the Leaked Facial Expressions: The Duration of Micro-Expressions', *Journal of Nonverbal Behavior* 37: 217–230.
- Zebrowitz, L. A., Voinescu, L. and Collins, M. A. 1996. '"Wide-Eyed" and "Crooked-Faced": Determinants of Perceived and Real Honesty across the Life Span', *Personality and Social Psychology Bulletin* 22: 1258–1269.
- Zhang, J., Li, X., Song, Y. and Liu, J. 2012. 'The Fusiform Face Area Is Engaged in Holistic, Not-Parts Based, Representation of Faces', *PLoS ONE* 7: e40390.

PART III

Legal Perspectives

11. Regulating the Expression of Remorse and the Building of Moral Communities

Richard Weisman

11.1 Introduction

This chapter builds on earlier research that the present writer and others have conducted over the past fifteen years into the critical role that expressions and attributions of different emotions play in legal decision making in general and in criminal law in particular. That law, despite its frequent portrayal as free of distracting passions, as nevertheless an instrument for the communication of these passions, is a conundrum that has now been explored with reference to shame, disgust, anger and other powerful emotions.¹ My research has contributed to this growing corpus by looking at the importance attached to expressions of remorse in decisions about sentencing and parole in Canada and the United States, and how the crediting or discrediting of these expressions creates a moral hierarchy of those who are remorseful and those who are without remorse that has far-reaching consequences for those transgressors so defined (Weisman 2009: 47–69). Perhaps the most graphic demonstration of the gravity and weight attached to these attributions stems from the work undertaken by the prolific National Capital Juror study in the United States. Numerous articles emanating from this project have attested to the importance jurors accord to the offender's show of remorse in deciding whether or not to impose the death penalty. In one of these studies involving California jurors in thirty-seven capital cases, it was found that 69 per cent of jurors who voted for death in the bifurcated penalty phase of the capital trial cited the absence of remorse by the defendant as the principal factor in their decision (Sundby 1998: 560).²

¹ For example, among the major works of the past fifteen years that have looked at the mobilization of shame and disgust in the law, see Nussbaum (2004); Bandes (1999); Kahan (2006); and, more generally, Miller (1998).

² See also the other most frequently cited article from the study on this topic, Eisenberg et al. (1998).

As of this writing, there is sufficient scholarship worldwide to suggest that similar concerns about remorse in which it is not just the transgressive act but the transgressor's feelings about the act play a major role in such disparate legal cultures as that of China, Japan, Sweden and the Netherlands, as well as in the common law jurisdictions of Canada, United States, England and Australia.³

Although my work has participated in this surge of scholarly interest into how attributions of remorse figure into judicial decision making, it has also taken a different turn of which this chapter is an outgrowth. Where much of the contemporary research has focused on the weight assigned to expressions of remorse as well as the arguments for and against allowing remorse as a mitigating factor in sentencing,⁴ I have advanced the claim that law plays a more fundamental role in the expression of remorse. Drawing on the pioneering work of Arlie Hochschild, I argue that law is a primary source for the regulation of emotions through the creation of 'feeling rules' (Hochschild 2001: 99–103) that instruct members of a community when they are expected to show remorse and how they are expected to show remorse (Weisman 2014: 2–3, 36). Because law is public and because it carries the authority of the state, the pronouncements of a court that a wrongdoer did not show remorse when they should have, or that their expressions of remorse were credible or not credible, has an impact on the community as a whole. It is this regulative role of law that I have expanded upon in this chapter, in which I have identified extra-legal as well as legal sources of social and moral regulation (for a related discussion of the role of emotions including remorse in social and moral regulation, see Helm, this volume).

I want to begin my discussion with two events from my time as a visiting professor at Hebrew University in Jerusalem in the fall of 2011. I have chosen them not to plunge into the quagmire of Middle East politics, but because they serve as useful points of reference to build on an approach to expressions of remorse that I have discussed in previously published work. The first event occurred three days after I arrived on

³ See especially Chang (2001); Johnson (2002); Dahlberg (2009); Komter (1998); Weisman (2009); Bandes (2014); Maslen (2015); Rossmannith (2015) for the discussion of the role of remorse in criminal justice systems in China, Japan, Sweden, the Netherlands, Canada, United States, England and Australia, respectively.

⁴ For one of the most thorough analyses of how remorse is used as a factor in sentencing as well as the different arguments for justifying this usage from the vantage points of jurisprudence, moral philosophy and psychology, see Proeve and Tudor (2010).

October 27, 2011. After several years of tense negotiations between Israel and Hamas, a number of Palestinians who had been arrested as terrorists were exchanged for a young Israeli soldier who had been abducted five years earlier by Hamas militants. But what made the local headlines that day and what the Israelis I met that day were discussing was not Gilad Shalit's feelings about his first day of freedom, but a televised interview with one of the newly released Palestinian prisoners. Ahlam Tamimi had been incarcerated since 2002 for her participation in what came to be known in Israel as the Sbarro massacre in which fifteen persons died as a result of a suicide bombing that took place in the Sbarro pizza restaurant in Jerusalem in 2001. Included among the fatalities were seven children ranging in age from two to eighteen years. When she was asked about her reaction to the fatalities, she was described as 'smiling broadly' and expressing satisfaction that fifteen persons including eight children had died instead of just three persons as initially reported.⁵ When asked how she felt about what she had done, she replied – 'Do you want me to denounce what I did? That's out of the question. I would do it again today, and in the same manner.' My hosts at Hebrew University would later use the word 'chilling' to explain their reaction to the interview and local newspapers would comment on her lack of remorse and her complete indifference to the deaths and casualties that had resulted from the bombing.

The second vignette refers to an event that occurred later in my stay. An Israeli lawyer whom I met through a mutual friend kindly agreed to give me a glimpse of the workings of the Israeli criminal justice system by allowing me to sit in on proceedings involving juvenile offenders. One of the defendants was a teenager of about fifteen from a settlement in the West Bank who had been charged with a hate crime against a Palestinian. She had destroyed property belonging to the victim and for no other reason than that the victim was Palestinian. Video evidence that showed her purchasing the equipment that would be used to commit the crime had made it into a virtually airtight case for the prosecution. While her lawyer wanted her to plead guilty and to show remorse in order to mitigate the sentence, the leader of the settlement insisted that she not plead guilty and that she not show remorse perhaps out of belief that her actions were morally justified and to oppose the

⁵ The interview was subsequently made available on You-Tube at www.youtube.com/watch?v=Iq28f0VztYw, retrieved on 15 January 2015. The version that was seen on Israeli TV is at www.youtube.com/watch?v=wkAuvlQTzWU, retrieved on 12 January 2015.

authority of secular courts to override the authority of religious courts. At the time that I lost contact with the case, the juvenile offender was in the thrall of cross-pressures between her desire for a reduced sentence and her loyalty to the leader of her community. In the following pages, I want, first, to identify more precisely the social processes by which feelings of remorse and their expression are regulated using these events as illustrations; second, to show the operation of these social processes during periods of social and individual change; and, finally, to give some indication of the pressures imposed on individuals and groups to adhere to these regulative norms by showing what happens when they do not.

11.2 Remorse and the Construction of Moral Communities

To put these events in context, I want to draw on the valuable work of Nicholas Tavuchis who postulated in his formulation of the ritual of apology that the ‘apology speaks to something larger than a particular offense and works its magic by a kind of speech that cannot be contained or understood merely in terms of expediency or the desire to achieve reconciliation’ (Tavuchis 1991: 7). Tavuchis’ important insight was to recognize that rituals of inclusion such as the offering of an apology or the showing of remorse if successful accomplish more than just the instrumentality of restoring a relationship between a person who was wronged and the person who transgressed against them. Through participation in these rituals, the wrongdoer establishes that they are a member of the moral community – that is, the community which expects a show of remorse or an apology after one has engaged in a transgression. Not to perform these rituals is to demonstrate that one does not belong to the moral community, either through defiance as communicated in the first example, or perhaps by silence as might be communicated by a refusal to show remorse after a finding of guilt by a court.

But I would like to modify and extend Tavuchis’ insight about rituals of inclusion and moral communities in several ways. First, as I have argued in earlier publications (Weisman 2009, 2014), although there are important similarities between the offering of an apology and the showing of remorse, there are also enough differences in usage and substance to treat the showing of remorse as a separate ritual that occupies a place in the building of moral communities quite distinct from that of the offering of apologies. If, following the work of Erving Goffman

(Goffman 1972: 113), we can conceptualize the social practices of both forms of communication as involving a splitting of the self between the person who breaks a rule and the person who condemns the breach, the two differ significantly in how this divided self is presented. The question that can always be asked with either moral performance is who is the real self – the person who committed the transgressive act, in which case the wrongdoer can be defined by what he or she has done, or the person who has united with the community in self-condemnation of the act, in which case the wrongdoer is defined by how he or she feels about what they have done. Where the making of the apology and the showing of remorse differ is in how this split between being and doing is communicated and how this communication is perceived. Elsewhere, I have argued that, with expressions of remorse, more than with the offering of apologies, the audience, whether it is a court or other institution or simply the victim, is likely to pay as much attention to the gestures, displays of affect and other paralinguistic cues as to the words themselves. I have further suggested that this distinction is rooted in the epistemological privileging of feelings over words in terms of revealing the core identity of the wrongdoer (Weisman 1999: 121–135). Hence, it is not accidental that so many criminal justice systems such as those in the United States, Canada and England, and in this example cited, Israel, dwell on expressions of remorse rather than the apology as the truer indicator of the authentic self. This emphasis on the non-verbal component of remorse is also evident in those who are judged as lacking in remorse. Just as the flooding out of uncontrolled emotions can be read as a sign of remorse, so also can the absence of any physical or psychological discomfort be read as the absence of remorse. The newspaper reports as well as the informal commentaries I heard on Ms. Tamimi's performance were all careful to note that she was smiling when she spoke of her participation in the bombings and that this corroborated that, not only did she say that she was not remorseful for what she had done, she showed through her bodily and facial expressions that she felt no remorse.

My second gloss on Tavuchis' work is to carry the concept of the moral community a step further. I want to suggest that participation in the ritual of remorse does more than qualify the wrongdoer for membership in a moral community. The very process of deciding whether an action should call forth a show of remorse and how remorse should be demonstrated to be credible is itself constitutive of the moral universe to which members of a moral community give their allegiance.

As we observe in the examples mentioned, what is in contention is not just whether the wrongdoer is remorseful or not, but whether these are actions for which one should show remorse. When groups are engaged in social and political conflict, one of the dimensions of that conflict is which actions call for remorse and which actions do not. And once we recognize the coexistence of multiple moral communities, it becomes possible to see how giving allegiance to one can be viewed as an act of betrayal by the other.

This, in turn, leads to another emendation of Tavuchis's rich concept. As we can see from the examples, moral communities exert pressure on their members both to show and not to show remorse. The pressures imposed by the law, which for present purposes I will conceptualize as the moral community as imagined by the state, are straightforward: not to show remorse when it is expected is to forgo the benefits of mitigation and thereby to risk a more severe sentence. But there are pressures as well not to show remorse when the community decides that this is not an act that should be accompanied by remorse. Ms. Tamini in the same interview describes the bus ride back to the West Bank in which the passengers were congratulating each other on the bombing and expressing joy in its success. The youthful offender in the second example was given clear instructions not to show remorse for an act that was contrary to Israeli law. Demonstrating remorse in either of these two circumstances would place the transgressor in conflict with the communities in which they define themselves as members.

11.3 Making Moral Communities Visible

But this work of moral communities in regulating the expression of the moral emotions of their members is rendered all the more difficult to discern because of the intensity and apparent spontaneity with which members react when their expectations for the showing of remorse are so egregiously disappointed. What appear to be flagrant violations of the moral code of a community arouse moral outrage not just because of the desecrating act but even more so when accompanied by defiance or indifference to the 'feeling rules' (Hochschild 2001: 99–103) or expectations that the community imposes on its members. The reaction to Ms. Tamini in the Israeli press and in local commentary – her rejection of the feeling rules even more than her involvement in the bombing – was passionate in its denunciation. From the vantage point of this community, only an extreme outsider could fail to feel remorse

and express remorse after committing an act that resulted in the deaths of so many children. So fraught and so powerful are these reactions to those whose refusal or inability to show remorse is the antithesis of what members define as a normal reaction that the work of the moral community in achieving this consensus appears as natural rather than as socially constituted. Hence, to make this work visible, it is necessary to look at communities or individuals in transition where prescriptions for – as well as prohibitions against – the expression of remorse are more likely to be explicit.

11.3.1 The Undoing and Rebuilding of Moral Communities in Wartime

The imposition of controls that forbid or inhibit the expression of remorse are likely to be most explicit when persons are recruited or conscripted into institutions or groups whose purposes and practices are antithetical to the groups with which they were previously affiliated.⁶ Studies and descriptions of military training for war offer rich illustrations of the social processes by which new recruits learn to overcome moral inhibitions towards killing. Indeed, some of the most famous speeches during World War II may be taken as applauding as heroic the very acts that in civilian contexts would have called for public expressions of remorse. General George Patton's memorable speech urging the killing of German captives during the invasion of Italy in 1943 vividly exemplifies this valorization of previously proscribed actions:

When we meet the enemy, we will kill him. We will show him no mercy. He has killed thousands of your comrades, and he must die. If you company officers in leading your men against the enemy find him shooting at you, and, when you get within two hundred yards of him, and he wishes to surrender, oh No! That bastard will die! You must kill him. Stick him between the third and fourth ribs. You will tell your men that. They must have the killer instinct. Tell them to stick him. He can do no good then. Stick them in the liver. We will get the names of the killer and killers are immortal. When word reaches him that he is being faced with a killer battalion, a killer outfit, he will fight less. (Weingartner 1989: 30)

However brief this powerful speech, we can already identify in rough outline the ideological reframing by which what was a moral taboo is transformed into a moral imperative. Here we find the invoking of themes of retribution, self-preservation and dehumanization of the

⁶ Examples from [Sections 11.3.1](#), [11.3.2](#) and [11.3.3](#) are a revised and updated version of Weisman (2014: 13–16).

enemy that enable members to shift between moral communities so that actions for which the community would earlier have demanded extreme contrition – killing someone who is helpless and vulnerable and who is pleading for mercy – are now viewed as essential for survival and the protection of community.

But moral communities need to be sustained on an ongoing basis and not only by moral exhortation. As Joanna Bourke makes clear in her enlightening history of the emotional impact of ‘intimate killing’ in World Wars I and II and during the Vietnam War, unwanted expressions of remorse over killing were feared because they could undermine morale as well as inhibit the aggression that was expected from soldiers. Here it was the absence of feelings of remorse rather than their expression that was normalized, even if the victims were innocent civilians – as recounted by a military chaplain who served in Vietnam – ‘if one has blown away innocent women and children but experiences no unpleasant feelings about it, then there is no issue left to resolve’ (Bourke 1999: 228). In case these feelings did not subside, they were pathologized by psychiatrists as abnormal, unmanly or ‘feminine’ or ‘neurasthenic’, implying a tendency to use illness as a sign of weakness or psychological ineffectuality suggesting a correlation between feeling and expressing remorse and a lack of character (Bourke 1999: 242–249). Lt. William Calley’s response to the charges brought against him testifies to the success of these prohibitions against feeling or expressing remorse. In March, 1971, after a lengthy and highly publicized trial, Calley was convicted of the premeditated murder of twenty-two infants, children, women and old men in what came to be known as the My Lai massacre during the war with Vietnam. The convictions represented only a small portion of the hundreds of killings for which Calley was originally charged with murder. The court imposed a sentence of life imprisonment which was commuted to house arrest by President Nixon one day later.⁷ Calley’s own response when he was first accused was one of disbelief and bewilderment.

I couldn’t understand it. I kept thinking though. Could it be that I did something wrong? I knew that war’s wrong. Killing’s wrong: I realized that. I had gone to war though. I had killed but I knew so did a million others. I sat there and I couldn’t find the key. I pictured the people of My

⁷ After various appeals, Calley would serve three and a half years of house arrest before his release. See Anderson (1998). For interviews with Calley and other persons charged with war crimes related to My Lai, see PBS documentary on the American Experience at www.pbs.org/wgbh/amex/vietnam/trenches/my_lai.html, retrieved on 18 March 2015.

Lai: the bodies and they didn't bother me. I had found, I had closed with, I had destroyed the VC: the mission that day. I thought, it couldn't be wrong or I'd have remorse about it. (Bourke 1999: 159)

Just in case such feelings should persist after discharge from the army as was observed among some of the veterans of Vietnam, another category could be applied – that of survivor's guilt – to suggest once again that the remorse that they felt was inappropriate, excessive and unwarranted (Bourke 1999: 238). Indeed, it would take several more years and hundreds of hours of 'rap sessions' during the late 1960s for these veterans to come together a month after Calley's conviction to collectively assert their right to feel and express remorse. In a ceremony held in Washington, DC, a month after Calley was convicted, more than eight hundred veterans gathered together to discard their medals, ribbons and other war mementos on the steps of the capital to show remorse for actions taken against the Vietnamese that they had come to abhor (Shatan 1973). In an influential article that appeared in the *New York Times* a year later, psychoanalyst Chaim Shatan described what he called post-Vietnam syndrome, only in this case he argued that acknowledgement of moral guilt rather than its suppression was the solution to the psychic damage wrought by war (Shatan 1972: 26). The abruptness of war is matched by the abruptness of war's end. The moral boundaries that were drawn during war are redrawn in the aftermath of war. Thirty years after My Lai, when Calley made a public statement that 'there is not a day that goes by that I do not feel remorse for what happened that day at My Lai', it would be before an audience that by now had come to expect it.⁸

11.3.2 *Regime Change and Remorse*

Other moments of visible and explicit intervention into the moral foundation of society are observable as well in the efforts of new regimes as they seek to break with past practices. The attempt by the Nazis to transform the moral sentiments of German citizens towards mentally and physically disabled persons provides another striking illustration of how loyalty to community comes to be measured as much by lack of remorse as by its obverse. Beginning in 1933, school children would be taken to asylums for the mentally ill, those who were neurologically impaired, and those who were severely physically disabled, there

⁸ Editorial, *New York Times*, August 28, 2009: A22. For discussion of winter soldiers, as this group of Vietnam veterans came to be called, see Morris (2015: 137–156).

to gaze at the display of human misery because in the words of one instructor, 'Every young person, every ethnic comrade of either sex who is interested in marrying must be led at least once through the screaming and nameless misery of an insane asylum, an institution for idiots, a residence for the crippled... Here he shall learn to appreciate the sacred genetic inheritance he has received' (Koonz 2003: 154). The purpose of these visits was to shift public perception from viewing the patients as sick persons who should be cared for by the community towards seeing them instead as objects of revulsion who were draining resources away from those who were healthy and could contribute to the new Germany. Here remorse over the mistreatment or neglect of the sick would be replaced by disgust over their physical imperfections and resentment at the burden that they placed on others.

In the later part of the twentieth and early twenty-first centuries, these radical shifts in political formation and moral architecture have taken a different form, although the undoing of one moral community and its replacement by another continues to play a central role in the transformation. Consider the follow exchange that took place during the hearings of South Africa's Truth and Reconciliation Commission in Pretoria, South Africa on May 24, 1999:

HATTINGH: In conclusion Mr De Kock, you have testified before various Amnesty Committees and during your criminal trial you testified in mitigation and you appeared before the TRC and during all these times, you expressed your sentiments about the deeds in which you were involved, the people who you killed, the people who you injured, the people whose property you damaged and so forth, in retrospect how do you feel about it?

DE KOCK: Chairperson, we wasted the lives of many people, not only those who we then regarded as the enemy, we also wasted the lives of our own people, young National Servicemen of 17 and 18 years of age, we destroyed young men of the same age in the ranks of the ANC and the PAC. There were cases in which people were not handling weapons personally or carrying weapons personally and during such incidents, we destroyed people completely, not only changed their lives, but destroyed their lives. We ruined the lives of their families in a sense, and changed their lives irrevocably. I feel that in all aspects, by living past one another, we destroyed one another for absolutely no purpose ultimately. It was a futile exercise... We obtained absolutely nothing as a result thereof. There will always be a yearning and a sorrow which will never disappear and which will never be able to be rectified. That is all.⁹

⁹ TRC Website: Amnesty Hearing, Day 1, Pretoria, 24 May 1999; see Official TRC website at www.justice.gov.za/trc, retrieved on 8 January 2015.

Less than three years before, on October 30, 1996, de Kock had been sentenced to two life sentences and 212 years imprisonment for what a justice of the South African Supreme Court had described as ‘revolting acts planned and executed in cold blood committed, condoned, and covered up by a system rotten to the core’ (Pauw 1997: 135). But these were actions for which de Kock as head of operations at Vlakplass, the secret police organization committed to the protection of the regime in power, had received honours and awards only a decade before.¹⁰ Now in a speech that would be endlessly scrutinized, de Kock appeared to be offering a public recantation in which he repudiated the political order that he had so recently served and instead professed allegiance to a newly emergent political order.

De Kock’s expression of remorse was in response to a question he had arranged to be asked by his lawyer. But raising the question of how those who had worked to uphold the racially based policies of the former regime under apartheid felt about their actions soon became the moment during the hearings that most captured the attention of the public and would become the focus of newspaper and televised coverage of the proceedings.¹¹ Here, those who appeared before the tribunal as applicants for amnesty would be judged over whether they expressed remorse, and, if so, whether their remorse was credible and sufficient to justify their inclusion within the new moral community. In de Kock’s case, the debate over whether he was remorseful would persist even after he was granted parole on January 30, 2015 (Bilefsky 2015: A10). Before his release, de Kock would say that he owed his survival to the intervention of black prisoners – as a convert from one moral community to another, he was quoted as saying that ‘the only friends I have now are my former enemies’.

¹⁰ The earliest public disclosure of the existence of Vlakplaas in 1994 in the final report of the Goldstone Commission of 1994 is described in the *Truth and Reconciliation Commission of South Africa Report*, Vol.2, Chapter 7, ‘Political Violence in the Era of Negotiations and Transition, 1990–1994’, Cape Town: The Truth and Reconciliation Commission, 1998. Eugene de Kock has given a brief account of the commission in de Kock (1998: 244–246).

¹¹ As Deborah Posel observes of the amnesty hearings, ‘their public function was to perform the catharsis of the apology, disclosure of wrongdoing, culminating in the recognition of, and apology for, the suffering inflicted’ (Posel 2009: 138). These exchanges would then be broadcast by the South African Broadcasting Association in its weekly highlights or emphasized in the documentary *Long Night’s Journey into Day*, the most widely distributed film on the TRC and nominated for an Oscar for best documentary of 2001.

Such confessional performances,¹² as political scientist Leigh Payne has referred to them (Payne 2008), are characteristic of societies undergoing transitional justice, in which the former regime is viewed as having engaged in state-authorized violence, and the new regime has promised a decisive break with this history. De Kock, as director of operations for Vlakplaas, occupied one of the most politically sensitive yet vital positions for the protection of apartheid in the face of growing opposition both within and outside South Africa. Through his demonstrations of remorse, he performed the anguished and painful transition from chief technician at Vlakplaas to aspiring member of the new society.¹³ His inward journey paralleled the changes that were occurring outside. The very actions that before had conferred pride and self-respect now became sources of shame and self-loathing. It was almost as if he were enacting on his person the transition from the maximally racialized society under apartheid to the multiracial democracy as contemplated under the new constitution. Participation in policies that advantaged one racialized group over another would henceforth be acts for which one should feel and express remorse.

11.3.3 Occupations as Moral Communities

However, we need not dwell on the extremes of war and revolutionary change to observe the collective work involved in the reshaping and monitoring of the moral emotions, learning when remorse is appropriate and when it is not is a feature of all moral communities. Even if these expectations are unwritten, inchoate and unenforced by any statute or diagnostic manual, this does not mean that their communication through approval and disapproval, acceptance and rejection, inclusion and avoidance are any the less effective as mediums of social regulation.

Jennifer Hunt's ethnographic account of female cadets joining an urban police force in the northern United States captures well both the informality of the process by which neophytes learn of these expectations and their transformative impact nonetheless. In one riveting encounter, a female rookie responded to a complaint by a pedestrian that 'something funny was going in the drugstore' (Hunt 1985: 321).

¹² For more on the idea that expressions of emotion in legal contexts may constitute performances, see Sorial, this volume.

¹³ De Kock's transformation is described in Gobodo-Madikizela (2003).

As Hunt describes the encounter, 'the officer walked into the pharmacy where she found an armed man committing a robbery'. When he pointed his gun at her, she 'still pulled out her gun and pointed it at him'. He ordered her to drop her gun telling her that 'his partner was behind her with a revolver at her head'. Again, she refused to lower her gun and the armed man, and the rookie remained in a stalemate until another officer entered the premises and 'ordered the suspect to drop his gun'. Initially, she believed that she had acted appropriately in the situation in which she had stood up to the assailant and the confrontation had ended with no loss of life. The officers in the precinct, however, believed she had acted in a cowardly manner by not firing her weapon. Hunt reports that a few months later, the rookie vehemently expressed the wish that she had killed the suspect and vowed that next time she would 'shoot first and ask questions later' (Hunt 1985: 322).¹⁴ Other equally powerful vignettes convey the process by which the female cadet learns to overcome her initial reluctance to use lethal force and to view as normal and even courageous responses that earlier would have been accompanied by expressions of remorse.

It is through such processes that the expectations for membership in a moral community are communicated whether or not it is the suppression or expression of the moral emotions that becomes the standard by which loyalty to the community is demonstrated.¹⁵ Occasions for the demonstration of this loyalty arise when actions that are valorized or merely normalized in one group are condemned by another. It is these moments that both define and reinforce the boundaries within and between groups for it is here that the refusal to show remorse becomes simultaneously an affirmation of membership in one moral community and a measure of estrangement from the other. For a member to show remorse for actions that the community has deemed honourable or principled or courageous is tantamount to betrayal just as the refusal to show remorse for actions that are viewed as heinous are occasions for collective moral outrage.

If it is appreciated that demonstrations of remorse or its absence refer not just to the relationship between the individual and the victim but also to the relationship between the individual and the moral community,

¹⁴ For further elaboration on the requirements for membership in this moral community, see Hunt and Manning (1991).

¹⁵ For another illustration of occupational socialization in which members learn not to apologize or express remorse for actions that draw censure outside their moral community, see Bandes (2005–2006: 475–494).

we can begin to understand why such moments can become the object of intense public scrutiny as well as occasions for the outpouring of passionate public responses. At stake in the decision of a court or a parole board or a tribunal to expect remorse or not to expect remorse is the choice between rival moral communities. Which moral community will be affirmed in such a contest, the community that condemns the act or the community that applauds the act? It is these stark choices that are being weighed in the examples with which I began this chapter.

11.4 Remorse, Moral Communities and Dehumanization

Deciding which actions should be accompanied by expressions of remorse and which should not forms only one of the coordinates by which moral communities are constructed. Equally importantly, moral communities distinguish among persons as well as acts by separating those whose suffering matters and to whom remorse is owed from those whose suffering does not matter and to whom we do not owe remorse, as we saw in the examples earlier regarding military conflict and revolutionary change with Nazi policy attempting to move the extremely disabled into the category of those whose suffering does not matter, and the Truth and Reconciliation Commission seeking to include the victims of apartheid within those groups whose suffering does matter.

Perhaps it is ironic then that the attribution of remorsefulness and remorselessness to individuals and groups is itself so frequently a filter for creating this moral hierarchy among transgressors. As I write this conclusion, the closely watched trial of Dzhokhar Tsarnaev is about to enter the penalty phase of his capital trial held in Boston in which a jury who has already found him guilty of all charges in connection with the deaths and injuries that occurred in the Boston marathon of 2013 will now decide between execution and life without parole. Foremost among the arguments, we are told, that the prosecution will use to persuade the jury to vote for death over life is that Tsarnaev has never expressed any remorse for what he did (Seelye 2015: A1). On this basis, arguably, from the vantage point of the prosecution, because the accused does not feel about what he has done the way he should feel, he has become one of the 'them' towards whom the community need not show mercy or forgiveness. Tsarnaev's putative lack of remorse for his crime obliges us to have no remorse for what may befall him. In this respect, the prosecution is engaging in a form of argumentation to make a defendant in a capital trial death worthy that has been written

about by many researchers including the present writer (Weisman 2008; Costanzo and Peterson 1994).

Just before I left Israel, one of my erstwhile colleagues pulled me aside after listening to a presentation I had given on my work on remorse to point out that the emphasis given to Ahlam Tamimi in the local press made it appear as if all Palestinians felt the way she did about the bombing. Just as do individuals so also do groups who fail to show remorse for what is perceived as a heinous crime become included in the category of those whose suffering does not matter and towards whom we need not feel remorse.

References

- Anderson, D. 1998. *Facing My Lai: Moving beyond the Massacre*. Kansas City: University Press of Kansas.
- Bandes, S. 2005–2006. 'Loyalty to One's Convictions: The Prosecutor and Tunnel Vision', *Howard Law Journal* 49: 475–494.
2014. 'Remorse, Demeanor, and the Consequences of Misinterpretation', *Journal of Law, Religion, and the State* 3:170–199.
- (ed.) 1999. *The Passions of the Law*. New York: New York University Press.
- Bilefsky, D. 2015. 'South Africa Grants Parole to a Death-Squad Leader', *New York Times*, 31 January A10.
- Bourke, J. 1999. *An Intimate History of Killing: Face-to-Face in 20th Century Warfare*. New York: Basic Books.
- Chang, Y. 2001. 'Culture and Communication: An Ethnographic Study of Chinese Courtroom Communications'. PhD Thesis: University of Iowa.
- Costanzo, M. and Peterson, J. 1994. 'Attorney Persuasion in the Capital Penalty Phase: A Content Analysis of Closing Arguments', *Journal of Social Issues* 50: 125–147.
- Dahlberg, L. 2009. 'Emotional Tropes in the Courtroom: On Representation of Affect and Emotion in Legal Court Proceedings', *Nordic Theatre Studies* 21: 129–152.
- De Kock, E. 1998. *A Long Night's Damage*. Saxonworld, South Africa: Contra Press.
- Eisenberg, T., Garvey, S. P. and Wells, M. T. 1998. 'But Was He Sorry? The Role of Remorse in Capital Sentencing', *Cornell Review* 83: 1599–1637.
- Gobodo-Madikizela, P. 2003. *A Human Being Died That Night: A South African Story of Forgiveness*. New York: Houghton- Mifflin.
- Goffman, E. 1971. *Relations in Public: Microstudies in Public Order*. New York: Basic Books.
- Hochschild, A. 2001. *The Commercialization of Intimate Life: Notes from Home and Work*. Berkeley: University of California Press.
- Hunt, J. 1985. 'Police Accounts of Normal Force', *Journal of Contemporary Ethnography* 13: 171–187.
- Hunt, J. and Manning, P. 1991. 'The Social Context of Police Lying', *Symbolic Interaction* 14: 51–70.

- Johnson, D. T. 2002. *The Japanese Way of Justice: Prosecuting Crime in Japan*. New York: Oxford University Press.
- Kahan, Dan M. 2006. 'What's Really Wrong with Shaming Sanctions'. Faculty Scholarship Series. Paper 102.
- Komter, M. 1998. *Dilemmas in the Courtroom: A Study of the Trials of Violent Crime in the Netherlands*. New Jersey: Erlbaum Associates.
- Koonz, C. 2003. *The Nazi Conscience*. Cambridge, MA: Harvard University Press.
- Maslen, H. 2015. *Remorse, Penal Theory, and Sentencing*. Oxford: Hart Publishing.
- Miller, W. I. 1998. *The Anatomy of Disgust*. Cambridge, MA: Harvard University Press.
- Morris, D. J. 2015. *The Evil Hours: A Biography of Post-Traumatic Stress Disorder*. New York: Houghton- Mifflin.
- Nussbaum, M. C. 2004. *Hiding from Humanity: Disgust, Shame, and the Law*. Princeton NJ: Princeton University Press.
- Pauw, J. 1997. *Into the Heart of Darkness: Confessions of Apartheid's Assassins*. Johannesburg: Jonathan Ball Publishers.
- Payne, L. 2008. *Unsettling Accounts: Neither Truth nor Reconciliation in Confessions of State Violence*. Durham, NC: Duke University Press.
- Posel, D. 2009. 'History as Confession: The Case of the South African Truth and Reconciliation Commission', *Popular Culture* 20: 119–141.
- Proeve, M. and Tudor, S. 2010. *Remorse: Psychological and Jurisprudential Perspectives*. Farnham, UK: Ashgate.
- Rossmannith, K. 2015. 'Affect and the Judicial Assessment of Offenders: Feeling and Judging Remorse', *Body and Society* 21: 167–193.
- Seelye, K. Q. 2015. 'Victims in Boston Marathon Bombings Are Split on Death Penalty', *New York Times*, April 21, A1.
- Shatan, C. F. 1972. 'Post-Vietnam Syndrome', *New York Times* 6 May: 26.
1973. 'The Grief of Soldiers, Vietnam Combat Veterans' Self-Help Movement', *The American Journal of Orthopsychiatry* 43: 640–653.
- Sundby, S. 1998. 'The Capital Jury and Absolution: The Intersection of Trial Strategy, Remorse, and the Death Penalty', *Cornell Law Review* 83: 1557–1598.
- Tavuchis, N. 1991. *Mea Culpa: A Sociology of Apology and Reconciliation*. Stanford, CA: Stanford University Press.
- Truth and Reconciliation Commission of South Africa Report*, Vol. 2, 1998. Cape Town: The Truth and Reconciliation Commission.
- Weingartner, J. 1989. 'Massacre at Biscari: Patton and an American War Crime', *The Historian* 1: 28–54.
- Weisman, R. 1999. 'Detecting Remorse and Its Absence in the Criminal Justice System', in Sarat and Ewick (eds.) *Studies in Law, Politics, and Society* 121–135.
2008. 'Remorse and Psychopathy at the Penalty Phase of the Capital Trial: How Psychiatry's View of "Moral Insanity" Helps Build the Case for Death', *Studies in Law, Politics, and Society* 41: 187–217.
2009. 'Being and Doing: The Judicial Use of Remorse to Construct Character and Community', *Social and Legal Studies* 18: 47–69.
2014. *Showing Remorse: Law and the Social Control of Emotion*. Surrey, UK: Ashgate Press.

12. Share Your Grief but Not Your Anger

Victims and the Expression of Emotion in Criminal Justice

Susan A. Bandes

12.1 Introduction

This chapter focuses on the growing trend towards using the criminal justice system as a venue for the expression of emotion by crime victims and their families. In the United States, Canada, the United Kingdom, Australia and New Zealand, crime victims and their families are being given the opportunity to address the court, to confront the defendant and to weigh in on how the case is being conducted. These innovations pose difficult questions for adversary systems, in which the trial has long been viewed as a two-way battle between the government and the accused.¹ In order to define and regulate the increasingly powerful role of the victim, it is first necessary to determine what goals the victim's participation is meant to achieve.

The legal context places the questions of this volume in sharp relief. Choices about who is permitted to express emotion in particular venues, and about what emotions may be expressed, have concrete consequences. In the criminal justice system, for example, the decision to afford victims a venue for expression has been linked to increased severity of sentence and increased likelihood of a death sentence. Ideally, a choice with such weighty consequences would be based on a considered judgement that permitting such expression will advance legitimate goals. Yet although the emotions of crime victims, and of murder victims' family members (hereinafter survivors), have been afforded an increasingly powerful role in the criminal justice system in the past several decades, the justifications offered for that increased role have been

¹ In inquisitorial systems, the role of the victim has generally been explicitly recognized and defined.

poorly articulated and inconsistent. Without a coherent understanding of what role the communication of victim emotion is meant to serve, it is difficult to determine what form the communication ought to take, and what limits should be placed on the communication. It is also difficult to determine, as an empirical matter, whether the communication is serving any legitimate purpose in practice, or whether it is in fact interfering with important criminal justice goals.

In the adversary system, the traditional view is that by representing the state, the prosecution adequately represents the interests of victims as well. Recently victims have been given increasing participation rights, but the rationale for providing that access has not been well articulated. At the same time, concerns have been raised about the effects of victim participation on values like fairness to the defendant, equality of treatment for defendants, the efficiency of the trial and equality of treatment for the victims themselves (Bandes 1996).² There are three possible rationales for providing a formal legal venue for victims to express their emotions. First, victims' emotions may provide relevant legal information that will aid judges, juries and other decision makers in determining guilt, sentence or other outcomes. Second, victims may find the opportunity to express their emotions therapeutic or cathartic. Third, victims' expressions of emotion may have a beneficial effect on the perpetrators, for example, eliciting empathy and remorse, which may, in turn, elicit forgiveness from the victim, thus positively impacting both victim and perpetrator.

Each of these rationales has its own implications for the nature, shape and significance of emotional expression. For example, if the expression is meant to be therapeutic or cathartic, this result may require some sort of response from the intended audience. It may require a judge to respond with affirmation or sympathy as opposed to silence or discomfort. It may require a jury to sentence the defendant to death, or a parole board to deny parole. Or perhaps it requires simply a forum

² I have written extensively and critically about victim impact statements and their impact on the criminal justice system. I have raised questions about their impact on the fairness of criminal trials in the capital punishment context (Bandes 1996), about the ways in which the prosecution agenda swamps the needs of victims (Bandes 1999), about their potential to exacerbate unequal treatment of minority defendants (Bandes 1996), about their potential to create classes of 'worthy' and 'unworthy victims' (Bandes 2000b; Bandes and Salerno 2014) and about their value as a therapeutic tool (Bandes 2009a). In this chapter, I focus on victim impact statements as a means of emotional expression, considering each of the articulated arguments for permitting the statements at criminal trials.

for expression, with no particular response necessary. Conversely, if the expression is meant solely to provide information to the decision maker, its effect on any particular victim should be irrelevant – the focus should be on what information is meant to be conveyed, and how to best convey it. Thus in a mass killing trial, for example, survivors should be permitted to speak only to the extent that their testimony provides information to the trier of fact, whether or not this decision deprives some survivors of an opportunity to be heard.

In this chapter, I will consider the criminal courtroom as a venue for the expression of the emotions of victims and their families. I will first examine what goals these emotional expressions are meant to serve, how these goals might affect the rules of expression and whether these goals are practically achievable or normatively desirable. Second, I will focus on the factors that shape victims' expression of emotion in the courtroom, and relatedly, on the factors shaping decision makers' evaluation and interpretation of victims' emotions.

The expression of emotion in the courtroom is not a recent development, of course. I have argued that emotion pervades the law (Bandes 2000a), and it certainly pervades the courtroom. Indeed, one of the most interesting questions about the expression of emotion in the courtroom is the question of what counts as 'emotion'. The expression of emotion in law is generally regarded with suspicion and disapproval, because 'emotion' is viewed as the opposite, or at least the enemy, of logic and rigor (Bandes 1996). The effect has often been to penalize the expression of *some* emotions, such as sympathy (*California v. Brown* 1986) and compassion (*DeShaney v. Winnebago County* 1989), while permitting or encouraging others, such as grief and some types of anger, that are so much a part of the fabric of law that their emotionality is invisible to the legal observer.

In short, victim impact statements, restorative justice conferences and other such innovations have not *introduced* emotion into the courtroom. I focus on these innovations because they recognize and formalize the expression of emotion. They thus provide a window into some of the criminal justice system's explicit attitudes towards emotional expression. Moreover, there is a growing literature in psychology, law and other fields on how victim impact statements and restorative justice conferences play out in practice. This literature sheds light on how the articulated assumptions about the expression of emotion compare to the unstated, implicit assumptions that also guide these encounters.

12.2 The Courtroom: An Emotion Culture

The criminal justice system is a fertile site for exploring the questions raised by this volume, both because it is unique and because it is emblematic. The courtroom is in many ways its own emotional ecosystem. It has explicit and implicit rules for the expression of emotion. Violating these rules can lead to serious consequences, including the loss of liberty and even (in the United States) to a sentence of death. More subtly, it also imposes feeling rules on its participants. Perpetrators are supposed to feel remorse, for example (see Weisman, this volume, for a discussion of feeling rules and remorse). But the courtroom context places in sharp relief the question of whether what is expressed must also be felt. In other words, how important is sincerity? In some realms, such as the spiritual or the therapeutic, sincerity may be of overriding importance. In the courtroom, emotion is always, in part, a performance with an end goal in sight (see Sorial, this volume, for development of this idea). Whether that performance must be sincere depends on the goals of the proceeding.³

There is no ignoring the fact that the courtroom is a place where judgement takes place, and where words, actions and other expressions of emotion are subject to that judgement. The dramaturgy of the situation is sometimes quite explicit. Attorneys are expected to perform, and it is a commonplace that their goal is to elicit emotion from jurors. More frequently, the drama is more subtle. For example, a defendant sitting silently in the dock listening to witnesses is giving an emotional performance, whether he chooses to or not. Jurors watch his demeanour and facial expression intently in the belief that they can intuit whether he is empathetic or remorseful (Weisman, this volume). Some of the drama occurs offstage (Rose and Diamond 2009); for example, in the emotional expressions and body language of family members sitting in the audience;⁴ and yet, it too becomes fodder for the ultimate judgement. Expression has concrete consequences in this context.

³ And if sincerity is important, that raises another difficult question: is the criminal justice system equipped to identify sincere emotion? Baker, Black and Porter, this volume, raise important concerns about our ability to determine whether or not emotional expressions are sincere.

⁴ See, for example, Rose and Diamond: 'Decision makers look not only to the behaviour of a given individual, but also how people behave *in relation to one another*, for example, whether a child sits near a parent, whether someone seems "intimidated" by a spouse, or whether someone is under the "domination" of a mother' (2009: 339).

Candace Clark (1997), Arlie Hochschild (1983) and others have illustrated the pervasiveness of implicit rules for the display of emotion in various contexts. Their work highlights the difficulty of cleanly distinguishing 'feeling' from 'display'. The fact that we may follow display rules does not necessarily mean we are insincere, or that performance is disconnected from feeling (Ekman, Friesen and Ancoli 1980: 1125). For example, as Clark showed in her examination of sympathy exchanges, we internalize a complex set of rules about who is eligible for sympathy, how much to express and how much we can expect in return. As she describes it, these exchanges are partially strategic; they are influenced by power differentials, and yet they may well be expressions of genuine sympathy. Hochschild memorably explained the notion of 'deep acting': in certain subjugated positions, it is not enough to go through the motions of following emotion scripts; one must also manage and ultimately internalize the correct feelings (Hochschild 1983: 113–114). Deviance from accepted or required expression norms elicits negative and even punitive reactions, and in order to avoid these negative consequences, one may hide one's feelings and put on a more acceptable display, or one may manage one's emotions to bring them in line with predominant feeling rules (Thoits 1989; also Baker et al. and Weisman, this volume). Thus the notion of pure, sincere emotion, unaffected by social norms and expectations about what to feel and how to show it, is unpersuasive and simplistic. Nevertheless, it is fair to say that in a courtroom, where the linkage between performance and consequences is so salient, the opportunities and incentives for the expression of sincere, unmediated emotion may be especially limited.

The courtroom illustrates the importance of approaching emotion from a sociological as well as a psychological perspective. In a lab, emotions are often approached as individual, internal phenomena. But in the courtroom, one can see the interplay of emotion and the ways in which the experience, expression and interpretation of emotion are shaped by social interchange. A witness testifies; jurors evaluate her credibility in light of their expectations for the display of emotion; jurors evaluate the defendant's demeanour as she listens to the witness; jurors evaluate the expressions of the victim, the victim's family and the family of the defendant; jurors watch one another as well as other legal actors in the courtroom (e.g., the attorneys) for cues on how to react to these displays; jurors deliberate and their deliberation affects their own emotions and their interpretation of the emotional performance of defendant, witnesses and other courtroom actors.

12.3 Victim Emotion in the Courtroom

Crime victims and survivors became a force to reckon with in the early 1980s. Because they lacked a formal role in common law criminal justice systems, they were often treated as an afterthought or entirely forgotten. Their most basic, uncontroversial set of goals was to be kept in the loop: to be notified of court dates and other developments in their cases. Other goals were more controversial. For example, victims sought increased input about the conduct of the trial – such as a chance to weigh in on the appropriate charge or sentence. This set of proposals was controversial because it would intrude on prosecutorial discretion (Bandes 1999). Victims sought the right to make statements at sentencing. These victim impact statements were controversial because of concerns they would violate the defendant's right to a fair trial (*Booth v. Maryland* 1987; Bandes 1996). Victims' rights initiatives in favour of the death penalty were controversial because they excluded the voices of survivors who did not support capital punishment (Acker and Karp 2006). Nevertheless, victims' groups achieved a number of their goals, most notably the right to deliver victim impact statements at trial, a right that is now firmly entrenched in common law countries. Victim impact statements are generally delivered at sentencing. They are not meant to serve as evidence of the crime. In upholding the constitutionality of victim impact statements at capital sentencing hearings, the US Supreme Court held that they are meant to inform the sentencing jury about the uniqueness of the victim and about how the loss of the victim impacted those left behind (*Payne v. Tennessee* 1991).

Although the Supreme Court upheld the statements on the rather narrow ground that they serve as a conduit for information, victim impact testimony ushered in a new era in which the criminal justice system offers victims and survivors a formal forum in which to express the pain and grief they experience as a result of crime. In short order the statements were recast, both in lower courts and in common parlance, as a way for victims and survivors to achieve healing, catharsis and closure (Bandes 2009a). Other formal opportunities for the expression of emotion have followed. In particular, restorative justice conferences, in which the victim is brought face to face with the offender in order to repair their relations, have become widely popular in Australia and New Zealand, and to a lesser extent, the United States and the United Kingdom. The emotional goals of the conference encompass both the victim and the

defendant. The defendant is meant to become a better person by gaining empathy and experiencing remorse. In turn, his understanding of the harm he has inflicted and his expression of remorse are meant to aid the victim's healing process. The victim may then forgive the defendant, which is thought to aid both parties (Acorn 2004).

The three rationales for providing a formal legal venue for the expression of victim emotion, as discussed, are sometimes explicitly articulated (e.g., the US Supreme Court has thus far upheld the use of victim impact statements under only the first – informational – rationale). More often, they are implicit, but no less powerful. The notion of the courtroom as a place for healing and closure has taken firm hold of the popular imagination (Zimring 2003: 60). It has also infiltrated the legal system, where it is often relied upon by lawmakers, by lower court judges (Bandes 2009a: 9–13) and by prosecutors exhorting capital juries to help victims heal by sentencing the defendant to death.

It is important to articulate and disentangle these rationales. Doing so is the first step towards investigating whether they have an empirical basis. In addition, the rationale dictates the shape of the victim's role. For example, if a victim impact statement is meant to provide catharsis for the victim, it is important to give every victim an opportunity to speak – even in a mass murder case. Alternatively, if the statement is meant to provide information to the court, it may be determined (as coldhearted as this may seem) that after a certain point, victim testimony becomes redundant, and thus (given its deeply upsetting nature) is more prejudicial than probative. If the victim's healing is the goal, the victim needs an unmediated opportunity to express sincere emotion (and whether the courtroom can ever provide such an opportunity will be discussed in the pages that follow). If, however, the goal is to evoke remorse in the defendant, or to publicly express the impact of the crime or to communicate the depth of the injury to the decision maker, it is not at all clear that the sincere, unmediated expression of emotion is of overriding importance. What becomes important in those situations is the ability to elicit the desired effect (this idea is developed by Sorial, this volume).

In the following section, I will consider the goals of permitting or encouraging victims to express emotion. Subsequently I will turn to the multiple ways in which a victim's emotional expression in the courtroom is influenced and shaped.

12.4 The Rationales for the Expression of Emotion by Crime Victims

12.4.1 *For the Victim: Healing, Closure and Catharsis*

The notion that the criminal justice system can provide crime victims and survivors with an opportunity to heal or attain catharsis⁵ has rapidly seized the popular imagination and become an article of faith for the media, the public and even courts and legislatures. Victim impact statements are often justified as opportunities for victim healing. Of the three rationales mentioned, this seems, at first blush, closest to the simplistic but commonly held view that emotional expression is an unmediated, sincere means of communicating one's internal emotional state (Ekman 1992). But even the most uncomplicated notion of expression must deal with a complication: expression rarely takes place in a vacuum. It cannot be understood or evaluated without consideration of the expected or likely response.

The victim impact statement is delivered by a crime victim, or – in a murder case – a member of the victim's family. A rape victim or the mother of a murdered child might turn to family, friends, a place of worship or a therapeutic professional to seek help with the healing process. In these venues, the victim's emotional expression will unfold as part of an interchange. Ideally the victim will have an opportunity to explain or expand on her feelings if needed, and even to communicate what sort of response would be most helpful. A courtroom victim impact statement is a speech (or at times a written document, read in court by legal personnel). It generally cannot be modulated in light of the reaction.

It is possible that healing or catharsis is promoted simply by the expression of feelings in an open courtroom, although there is thus far little empirical evidence to support this assumption (Armour and Umbreit 2006). This goal requires simply the provision of a venue for sincere expression. As I will discuss, the 'sincere expression' aspect of this goal is unlikely to be realized, since the expression is meant to serve other, sometimes conflicting, purposes as well. But if the pure

⁵ Healing and catharsis are not synonymous. The former suggests an internal, therapeutic process. The latter may also refer to internal processes, for example purgation and cleansing, but, as used in the dramatic arts (such as in the Aristotelian notion of tragedy), implies a more public process directed at evoking emotion from the audience. As to closure, I have argued that this is a highly imprecise umbrella term that encompasses a number of disparate concepts (Bandes 2009a: 16).

opportunity to speak is indeed the goal, that simplifies the equation, in the sense that the statement does not need to carry any evidentiary weight, and no legal decision needs to hinge upon it.

The notion of catharsis through expression does raise the question of whether the opportunity to speak in court is essential. There is evidence that victims whose written statements were delivered by legal personnel felt less satisfaction with the process than those who were permitted to speak in court in their own voices (Davis and Smith 1994: 11). A related question is whether a videotaped statement would suffice, or whether the victim's physical presence in the courtroom is integral to the process. Finally, the healing and catharsis frame raises the question of whether the victim's opportunity to speak must take place during trial, or whether some ancillary process could be established. It seems highly likely that an in-court audience during the actual trial is an important part of the process. The victim impact statement is considered part of the sentencing hearing (though its legal weight in sentencing is unclear and unresolved). The declaration of verdict and sentence serve an expressive function, announcing to the community the law's condemnation or exoneration of the defendant. A victim impact statement is different from a private expression of grief precisely because it occurs within this formal, official, public context.

Conversely, the victim impact statement is very much like an expression of grief in an unofficial (e.g., a familial or therapeutic) setting in the sense that, as in any social interchange, the victim or survivor delivering the statement hopes for or even expects a reaction of some sort. If so, the question becomes: which audience matters, and for what legal purpose? The legally recognized purpose of the statement is to provide the trier of fact with information, and thus the statement is formally addressed to the sentencing judge or jury. But the statement has additional implicit and explicit audiences. The expressive function of the statement (and of the sentencing hearing more generally) assumes a wider public audience. Indeed, victim counselling services explicitly remind those preparing victim impact testimony that the media may take an interest in their statements (Mothers Against Drunk Driving – MADD – 2013). And as an emotional matter, there is a sense in which these statements are addressed, ultimately, to the offender.

In a capital case, the jury, once it has convicted the defendant, determines whether a death sentence is appropriate. I will return to the question of what information the victim impact statement is meant

to provide to the jury. The immediate question is whether the jury is meant to contribute to the survivor's healing process. If so, the jury has limited options available. One is simply to show appropriate facial emotion during the delivery of the statement. But for a penalty-phase jury that wishes to act on its sympathy for the victim's family and its anger and disgust towards the defendant, the only action available is to reflect those emotions in the sentence. There is substantial evidence that these emotional responses in jurors are linked to an increased likelihood of a death sentence (Paternoster and Deise 2011).

Prosecutors routinely assure juries in their closing arguments that their decision to put the defendant to death will provide closure for the victim's family (as well as affirm their judgement of the victim's worth). One intriguing possibility is that survivors' emotions have been shaped by this newly minted but insistent message: that survivors now expect that the measure of the effectiveness of their statement describing the impact of the loss of the victim is whether it leads to a death sentence, and that only a death sentence can help them heal (Bandes 2009a).⁶ Another possibility, to be discussed in detail, is that punitive agendas belonging to others swamp the survivors' ability to express and experience their own pain (Arrigo and Williams 2003: 618–619).

Note the implications of this linkage between expression and audience for the notion of sincere, unmediated expression of emotion. First, sincere, unmediated emotion may not be best calculated to elicit a death sentence. It may not conform to the jury's expectations about the appropriate expression of grief (Rose, Nadler and Clark 2006), and for that or other reasons it may not elicit the anger and disgust towards the defendant, or the sympathy for the victims and survivors, that have been linked to an increased likelihood of a death sentence (Paternoster and Deise 2011: 641–642). Second, a survivor whose sincere emotions include a belief in redemption and forgiveness faces a quandary, having been told that a death sentence is a measure of the worth of the victim and that the failure to seek a death sentence implies a lack of respect for the victim. At bottom, the notion of sincere, unmediated emotion is problematic in this context. When expression has legal consequences,

⁶ To the extent that survivors do internalize this message, it brings another troubling possibility: that of secondary victimization when the hoped-for sense of healing and release does not materialize despite the defendant's sentence or execution (Acker and Karp 2006: 253).

one is likely to modulate one's expression in light of those consequences, or even to modulate the feelings themselves.⁷

In non-capital cases in the United States, the victim impact statement is delivered to the sentencing judge. The question of whether the statement is aimed at affecting the sentence applies in this context as well as the jury context. But there are important differences between the jury and the judge as audience. A judge is less constrained than a jury in her ability to react to the statement, at least in a formal sense, and this ability to react raises another set of questions about audience. First, it means the judge must explain her decision, and thus must confront the ambiguous evidentiary status of the victim impact statement. Is the statement simply an opportunity for the victim to speak, in which case the judge should accord it no weight in sentencing, or is it a factor in the calculus? If the former, the judge must negotiate a difficult task. She must listen to the highly emotional testimony without appearing to denigrate its seriousness, while giving it no weight in her decisional calculus (Schuster and Proven 2011: 59).

In most venues, it would be exceedingly strange to respond to an emotional statement about grievous suffering by offering a neutral affect and remaining entirely silent. Yet the courtroom is indeed its own unique culture in this regard. There is perhaps no script with greater staying power than the script that commands judicial dispassion (Maroney 2011). As mentioned, the legal system reflexively equates emotion with illogic. The jury is partially exempt from the ban on emotional expressiveness, because the jury is made up of laypeople representing the larger community. The judge, however, is regarded as the apotheosis of reason in a culture that equates emotion with unreason. Judges may believe that expressing sympathy towards an individual undermines objectivity or the appearance of objectivity. Even for a judge who questions these role expectations, there is the problem of training. Whereas a therapist, a member of the clergy or even a friend confronted with suffering will have some sort of script to draw from, judges have little guidance about how to show emotion in the courtroom⁸ without detracting from their aura of impartiality or from their authority – or

⁷ As noted, the pressure on survivors to support the death penalty may influence not only their stated positions about the relationship between capital punishment and healing, but also their own internal emotional expectations.

⁸ Schuster and Proven (2011: 85) report that some of the judges they interviewed solved this dilemma by addressing the victim after the end of formal proceedings and outside of the courtroom.

about which emotions are appropriate or helpful. One judge observed: 'I oftentimes don't know what to say' (Schuster and Proten 2011: 86–87). Others were afraid that giving in to emotion would break down the wall they needed to maintain between the pain they heard about every day and their own equanimity (Schuster and Proten 2011: 65). And thus it is common for judges to listen quietly to emotion-laden statements and to maintain a neutral affect.⁹

Victims who make themselves vulnerable in the intimidating atmosphere of a courtroom, unsurprisingly, tend to want more than neutral silence in response. Victims do better 'if the judge appears to listen to them, speaks to them directly, or even acknowledges their... [statement] when imposing the sentence' (Schuster and Proten 2011: 84; Erez and Laster 1999: 553). Some judges noted that they made the effort to look the victim in the eye rather than taking notes (Schuster and Proten 2011: 85). Some elected to address the victim. Validating or encouraging words from a judge can have a powerful impact on victims. Miller (2013) reports an interview with a rape victim whose victim impact statement moved the judge to assure her that she was not to blame for what had happened to her and to acknowledge the pain and suffering she had endured. The victim reported: 'Because of what the judge said, it was so easy just to walk out of that court and start my life' (Miller 2013). Schuster and Proten (2011: 86) reported that statements made to a defendant in the presence of the victim, such as 'I can't believe how much damage you have caused here', can be validating for the victim as well.

In short, the notion that the delivery of victim impact statements can help victims and survivors heal has seized the popular imagination, but is a highly problematic rationale for permitting the testimony. To the extent it demands a reaction from the jury, there is no good reason to believe that a punitive jury verdict will help victims heal. To the extent it demands a reaction from a judge, the judiciary lacks the training to react in ways that will promote healing – and indeed is substantially constrained by role expectations that discourage affective display.

12.4.2 *For the Defendant: Inculcating Remorse, Empathy or Shame*

The ambiguous communicative status of the victim impact statement is illustrated by the fact that for victims, the defendant may be the

⁹ In addition, judges at times exhibit reactions that are dismissive or patronizing towards victims or their accounts (Logan 2008: 769).

audience who matters most as an emotional matter. The MADD workbook, one of the most influential resources for victims, describes the victim impact statement as an opportunity to present 'the true realities of victimization to those who haven't gone through it themselves and to the offender who has hurt you'. Yet victims are often instructed to ask the court's permission if they wish to address the defendant directly, or to 'say what you need regarding the offender through the judge' (Victim Support Services 2015).

Victims sometimes report that the mere opportunity to face the defendant is healing. In the Boston Marathon bombing trial of Dzhokhar Tsarnaev, Rebekah Gregory, who lost a leg in the bombing, wrote an open letter to the defendant after giving testimony, telling him that even though he never looked her way: 'I realized that sitting across from you was somehow the crazy kind of step forward that I needed all along' (Seelye and Bidgood 2015). Others seek some sort of acknowledgement. One mother addressed her daughter's murderer directly, telling him she had 'no room in her heart for hating him. But the defendant would not look at her, nor at the photo she held of her daughter. He stared impassively ahead, his big sloping shoulders still as a rock.' She said later: 'I wanted to make sure he knew I was there.' But she 'could see nothing in his eyes' (McKinley 2014).

Confronting a defendant with the impact of his crime is sometimes justified as a means of positively influencing the defendant, which will in turn help the victim as well. Stephanos Bibas and Richard Bierschbach provide an example of the ideal script for this dynamic:

When criminal justice does produce remorse, the effects can be profound. When victims' relatives confronted serial killer Gary Leon Ridgway at sentencing, they sobbed and poured out their anger and loss. The judge expressed the community's moral condemnation and spoke of bringing peace and closure. In return, Ridgway expressed sorrow and apologized, and at least one victim's relative forgave him and expressed a feeling of peace. (Bibas and Bierschbach 2004: 87–88)

The hope for a satisfying response from the defendant may be central to many victims and survivors, but it is far from guaranteed. In capital cases, the defendant rarely testifies. A defendant sentenced by a judge has the opportunity to acknowledge responsibility and even express remorse,¹⁰

¹⁰ One US jurisdiction has just made it mandatory for defendants to be present for the delivery of victim impact statements. See http://sentencing.typepad.com/sentencing-law_and_policy/2015/08/new-hampshire-enacts-novel-law-requiring-defendants-presence-in-courtroom-for-victim-impact-statemen.html.

but whether he will do so in a satisfying way is another issue (Weisman, this volume) – and whether his apology will ring sincere when it is delivered as part of his plea for a lesser sentence is yet another issue. There is also the horrifying possibility that the defendant will express emotion, sincere or otherwise, which is upsetting or destructive to the victims.¹¹ In summary, the notion of the victim impact statement as a means of moral education, is the least developed of its possible rationales. It is a poor fit with the formal structure of the statement as a means of conveying information to the judge or jury. Nevertheless, for victims and survivors, the opportunity to address the defendant appears to be emotionally powerful.

An interactive, mediated process seems better suited to the goal of aiding the victim in communicating the impact of the crime directly to the defendant, and victim-offender mediation (also known as a restorative justice conference) is meant to accomplish precisely this goal. This is a ‘process whereby the parties with a stake in a particular offense come together to resolve collectively how to deal with the aftermath of the offense and its implications for the future’ (Rossner 2013: 8). The core elements of these encounters are: a defendant who accepts responsibility for the crime; an opportunity for everyone present to express their emotions; and an effort to reach general consensus about how the offender can redress the harm (Rossner 2013: 9). In short, the expression of emotion is front and centre in this venue. Unlike the victim impact statement, this encounter does allow for face-to-face interchange. By definition, the exchange is not unmediated. A facilitator screens and prepares the parties beforehand, and participates in the discussion. Whether this type of encounter allows for, or encourages, sincere interchange is a separate issue, which I will address in what follows.

12.4.3 *For the Court: Providing Information*

The actual legal basis for victim impact testimony is the most difficult rationale to comprehend, much less evaluate. According to the US Supreme Court (which has ruled on victim impact testimony only in the capital context), the victim impact statement is meant to provide

¹¹ For example, William Bonin, who raped, brutalized and murdered twenty-one boys and young men, delighted in torturing their parents in response to their requests for information about their sons’ deaths (Bandes 2000b: 1599).

information to the court. Specifically, it is meant to remind the sentencer that the 'victim is an individual whose death represents a unique loss to society and in particular to his family' (Payne: 825). At the same time, victim impact statements are not meant to encourage comparative evaluation of victims or to imply that some victims are worthier than others (Payne: 823). As Jessica Salerno and I have discussed in detail elsewhere, this is the puzzle at the heart of *Payne* (Bandes and Salerno 2014: 1030): if the goal is not comparative, what exactly *is* the testimony meant to express? How does one convey uniqueness in a way that is relevant to determining sentence without encouraging comparison with other victims? Moreover, to the extent that the victim's uniqueness, or the uniqueness of his loss, turns on factors irrelevant to sentencing, the testimony is problematic – and particularly problematic, especially if these factors include race, ethnicity, class, gender or other unconstitutional influences on judgement.

Despite the court's effort to cast victim impact testimony in purely informational terms, there is no question that it conveys emotion. Indeed, the emotional impact of the loss is difficult to separate from the court's justifications for the testimony: it ensures that the victim is not presented as a 'faceless stranger' and helps the State to convey 'the full moral force of the evidence' (Payne: 825). The actual experience of hearing the testimony is deeply emotional. One federal judge recalled: 'It has now been over four months since I heard this testimony...and the jurors' sobbing during the victim impact testimony still rings in my ears' (Bandes and Salerno 2014: 1032). If the goal of the statements is to communicate the impact of the loss, then these strong emotional reactions may be a testament to their effectiveness. As Salerno and I have suggested elsewhere, in murder cases in particular the statements may serve to elicit an 'identifiable victim effect', by transforming faceless, silenced victims into flesh and blood people by personalizing them and their families (Bandes and Salerno 2014: 1033–1034). Nevertheless, the question remains what the jury is meant to do with this information. The identifiable victim effect predicts that people will be more likely to care about and want to help identifiable victims than they will faceless, 'statistical' victims. One assumption embedded in the *Payne* reasoning is that in a capital case, a death sentence helps the victim's family,¹² and that for the testimony to be effective it must increase the likelihood of

¹² The court stated that victims and their families have 'something to be gained or lost from the jury's verdict.' *Payne*, 501 U.S. at 826.

a death sentence. But not every victim's or survivor's grief is accompanied by the desire for harsh punishment.

In summary, it is difficult to evaluate whether victim impact statements successfully convey relevant information, given the arguably fatal ambiguity about what information they are meant to convey, and to what end. But the salient point here is that if the goal is informational, then the reception of the statements by the fact-finder matters to the determination of whether that goal is achieved.

12.5 Display Rules: Communicating and Interpreting Victim Emotion

I have argued that emotional expression is governed by implicit and explicit rules of display, and that the courtroom both reflects the display rules of the society in which it is embedded and imposes its own set of display rules. I have also argued that whether an emotional expression is 'successful' depends on what it is meant to achieve, and that there is significant ambiguity about what the expression of victims' emotions in the courtroom is meant to achieve. Whether the goals are victim-centred (healing), defendant-centred (moral education) or decision-centred (information), the criminal justice system lays down implicit and explicit rules governing what a victim may express. Some of the factors shaping and channelling victim expression take the form of direct influences on the victims themselves. In addition, audience expectations shape victim expression. Deviance from these expectations increases the likelihood that the expression will be misinterpreted or will elicit the wrong reaction, and that the goals of the expression will not be achieved.

12.5.1 Explicit Rules and Expectations

The liaison between crime victims and the criminal justice system in the United States, and largely in other common law countries, is generally the prosecutor's office. Victims and survivors are state's witnesses, and when they prepare victim impact testimony, they do so in conjunction with the state or with social service agencies working with the state. Although the victims' rights movement has sought to carve out an independent role for victims, this structural set up remains the norm. It carries with it the assumption that the victim's interests are aligned with the state's interest. Although the state is formally meant to have

dual interests – obtaining justice and obtaining a conviction – the latter is usually the more salient (Bandes 2006). I have written about the issues arising from this assumed alignment between victim and prosecutor in detail elsewhere (Bandes 1996, 1999; Bandes and Salerno 2014) and cannot do them justice here. I can only briefly highlight some of the implications of this alignment for the expression of emotions by victims and survivors.

The most dramatic consequence of this has been the refusal of prosecutors to permit certain survivors to give victim impact testimony in capital cases. In the United States, the prosecution's refusal to permit victim impact testimony by the mother of a child murdered in the Oklahoma City bombing on the grounds that she opposed the death penalty helped spur the passage of federal legislation that sought to protect victims' right to give such testimony whatever their views. Nevertheless, for a variety of reasons (including lack of enforcement mechanisms for the federal legislation or for similar state-level protections), survivors who do not support a death sentence continue to experience difficulties giving victim impact testimony (Acker and Karp 2006: 253). A more subtle silencing of certain emotional views occurs as well. Victim impact witnesses in death penalty cases are often barred from opining on the appropriate sentence, yet certain emotional stances towards punishment are so deeply ingrained that they are not coded as emotional pleas in favour of a certain outcome; they are regarded as simply factual, affectless statements. Witnesses who mention compassion or mercy in their statements are often barred on the grounds that they are advocating against a death sentence, whereas witnesses who mention vengeance or advocate for bringing the full weight of the law to bear on the defendant are regarded as simply stating the law, rather than expressing an opinion on the appropriate sentence (Bandes 2009b: 496–498). The result of this differential treatment is not merely to bar anti-death penalty victim impact testimony – it is more pernicious than that. Because survivors are not permitted to opine on sentence, but are at the same time testifying on behalf of the state, the very act of giving victim impact testimony at a capital trial renders the testimony 'government testimony' and aligns the survivor with the prosecution's goal. Victim impact statements serve prosecutors' message that the death penalty should be imposed for the sake of the victims. Those who do not wish to be aligned with that message have few if any formal means to dissociate from it. This quandary was recently highlighted in the Boston Marathon bombing trial of Dzhokhar Tsarnaev.

The Richard family, whose loss of their eight-year-old son and grievous physical injury were a prominent part of the government's case, had no formal avenue for dissociating itself from the prosecution's demand for capital punishment, and thus the Richards wrote an open letter to the Justice Department – a letter that the government may or may not choose to act upon at its own discretion, but that the jury in Tsarnaev's case did not see before it began deliberating on his fate (Richards and Richards 2015).

The assumption that victims testify in order to elicit more punitive sanctions puts pressure on victims to express the emotions that are expected rather than those perceived as deviant. The message is quite clearly communicated that certain emotions are not appropriate and that expressing them will not lead to the desired results. In the courtroom, the desired result is generally assumed to be increased punishment. In the restorative justice context, which I will touch on briefly later, the desired results are the defendant's remorse and the victim's forgiveness. In the next section, I will discuss expected emotions and emotion deviance in the contexts of capital cases, rape cases, domestic violence cases and victim offender mediation.

12.5.2 Expected and Deviant Emotions: The Implicit Rules

Rules for the feeling and display of emotion exert a powerful effect on judgements of credibility and other judgements of character in the courtroom. When victims or survivors testify, they are subject to judgements about credibility that draw on accepted scripts about how victims 'ought to' feel and express those feelings. The prototypical victim becomes the 'measure against which judges and other legal professionals evaluate the truthfulness and credibility of victim accounts' and atypical victims 'risk being dismissed as exaggerated, illogical and noncredible' (Erez and Laster 1999: 542). Victims are evaluated not only for their credibility, but according to how sympathetic they appear, and the emotions they exhibit are part of this calculus as well. Victims are judged by whether they are exhibiting the appropriate emotions for the situation, and at the expected level of intensity. These role expectations, unsurprisingly, are affected by a number of variables, such as gender, race (Aguirre et al. 1999), social class and ethnicity, in various ways that cannot be adequately explored here. Instead, I will focus on one particular variable – gender – and some of its effects on the evaluation of emotional expression.

I earlier raised the question of what kinds of information a victim impact statement provides to the judge or jury at a sentencing hearing. There is a growing body of literature on how victim impact testimony impacts decision making. In discussing this literature, it is useful to distinguish capital murder trials, in which the victim's survivors are testifying at sentencing, from rape, domestic violence and other crimes that do not involve loss of life, where the victim herself is testifying. Murder cases present a particular conundrum, because the victim impact testimony is not directed at the circumstances of the crime or the character of the defendant, both of which are addressed in other testimony. It is directed at the uniqueness of the victim and the impact of the loss on the victim's family. Thus, despite the US Supreme Court's dictum that the statements are not meant to establish the comparative worthiness of victims or their families, these sorts of comparisons do in fact occur. Victim impact statements, in both capital and non-capital cases, generally increase juror anger towards defendants and empathy towards victims and their families, and they generally lead to increased sentences (Deise and Paternoster 2013: 615). Some studies have found, though, that when jurors viewed victims as respectable they felt more empathy for the victims, rated the harms as more serious and felt more compassion for the victim's family members (Greene, Koehring and Quiat 1998). There is evidence that these views are shaped by various factors that ought to be external to judgement, including race (Bowers, Sandys and Steiner 1988) and social class (Paternoster and Deise 2011: 154). They are strongly shaped by assumptions about gender. Sundby, for example, found that the murder of victims who had placed themselves in situations the jury perceived as dangerous or deviant (such as women drinking and picking up men in biker bars) was less likely to lead to a death sentence than the murder of victims doing what jurors could imagine themselves doing or approved of people doing (such as women withdrawing cash from an ATM machine) (Sundby 2003: 374–375).

Rape victims who give victim impact testimony are also judged by their character. Judges exhibit the same sadly persistent prejudices about 'real' rape that afflict laypeople, perceiving the statements more sympathetically in cases of stranger rape or when the crime is perceived as particularly egregious (such as the rape of a grandmother) and less sympathetically when the victim is perceived as complicit, for example, where the victim was dating or drinking with the offender (Miller 2013: 1447). One study found that prosecutors refrained from seeking victim

impact statements at all in cases where they did not regard the rape as egregious, or the victim as a true victim (this was a particularly prevalent reaction where the complainant was a prostitute). It further found that prosecutors felt hindered by victim impact statements that were 'unsophisticated' or 'colloquial' and thus failed to powerfully convey the full impact of the harm (Miller 2013: 1450).

Expectations about appropriate victim emotion take a number of forms. Some expectations concern the intensity of the emotion. Rose, Nadler and Clark (2006) posit a 'proportionality norm' where the intensity and duration of the emotion are expected to be proportional to the seriousness of the crime. They found evidence that these expectations about the intensity of emotion are gendered – that women are expected to react more intensely. Victims who fail to muster sufficient emotion in response to a crime are viewed as less credible. Most important, the emotions displayed must also be the right emotions. And one common finding that links various criminal justice contexts is that anger is not the right emotion.¹³

Victim impact statement workbooks generally encourage victims and survivors to speak from the heart (MADD 2013). Yet they frequently contain some variant of the following suggestion: 'Don't vent your anger toward the court or the offender. Your goal is to express your hurt and your pain, not to blame' (Victim Support Services 2015). Of course, if the victim's goal is truly to speak from the heart, the expression of anger and blame may be essential to attaining that goal. But if the victim's goal, or a goal of the prosecution for which the victim is being explicitly or implicitly conscripted, is to obtain a more punitive sentence, then forgoing the expression of anger appears to be wise counsel. Schuster and Proppen (2010, 2011) observed the delivery of a number of victim impact statements in rape and domestic violence cases in Minnesota (in which the audience for the statement was the presiding judge). They found a clear hierarchy among emotions, with compassion and the 'correct' amount of grief (Schuster and Proppen 2010: 90) viewed most favourably

¹³ Two related observations are in order. First, as mentioned, the desire to see the defendant harshly punished is not always coded as anger – it is sometimes viewed as simply rational. Second, there is evidence that anger may be perceived as more appropriate and less problematic when expressed by a man rather than a woman. See, for example, Barrett and Bliss-Moreau (2009). The Shuster and Proppen study discussed in what follows focuses on rape and domestic violence cases in which women comprise the vast majority of the victims. More work needs to be done on the role of gender expectations in shaping feeling and display rules in the courtroom.

as civilized and appropriate emotions, and anger viewed as a 'lower' and 'weak emotion. In rape cases, they found that judges viewed anger in the courtroom as 'unwelcome, uncomfortable and unproductive' (Schuster and Proten 2010: 92). Judges viewed a victim's anger as immature, as evincing a lack of perspective and as undermining the victim's credibility as a witness (Schuster and Proten 2011: 68). They also viewed it as a type of role confusion. In their view, moral outrage at the crime was the province of the court. Anger by the complainant connoted a thirst for revenge, which they did not view as the proper province of the victim. Compassion was the 'celebrated emotion' in the courtroom, requiring a higher level of insight and growth. Judges viewed compassionate victims as more credible than angry victims because they were able to 'transcend' the crime rather than let strong emotion interfere with their objectivity. Moreover, judges felt 'comfortable' and 'relieved' with compassionate victims, whom they saw as moving 'beyond a focus on their personal loss' and accepting 'the goals of the court', thereby acknowledging the authority of the judge and his control of his courtroom (Schuster and Proten 2011: 69–70).

Interestingly, victim-offender mediation, which is presented as a more authentic, informal opportunity for the expression and exchange of sincere emotion, has been critiqued for its enforcement of a similar hierarchy. Restorative justice conferences have their own set of scripts, and the goals appear to be focused mainly on remorse and forgiveness. Acorn argues that:

Restorative justice gives dramatic space to the victim's vengeful longings, while at the same time denying the force of those longings as reasons for either state or private action...The perpetrator is to be protected not only from the actual infliction of any retributive suffering, but also from the potentially degrading experience of being 'dumped on' by... the victim... [R]estorative justice further promises that this method will purge the victim's desire for vengeance and restore the victim to a sense of dignity and well-being. (Acorn 2004: 54–55)

Although this tendency to place anger low on the emotion hierarchy appears in a number of criminal justice contexts,¹⁴ Schuster and Proten (2011: 90) found that in the courtrooms they studied, it did not extend to domestic violence cases. Judges frequently viewed domestic violence victims as complicit in the violence they suffered, and viewed

¹⁴ See Jacoby (1983: 358; people tend to feel more sympathetic towards the forgiving victim than towards the angry, resentful victim).

their compassion as suspect and as undermining their credibility. The victim's compassion raises the bar for her credibility: she must now pass additional 'truth tests'. The compassionate domestic violence victim will be regarded as credible only if she is in the process of leaving her abuser or takes responsibility for her role or makes suggestions on how to protect herself. Otherwise she is granted little agency (Schuster and Proppen 2010: 100).

12.6 Conclusion

The courtroom has its own emotion rules and emotional dynamics, which make it unique in some regards. Context matters: the communicative role of emotion and its normative significance depend on the arena in which the emotions unfold. At the same time, the courtroom reflects the larger rules and assumptions of the society in which it is embedded. Certain global aspects of emotional expression are particularly salient in the courtroom: its inherent dramaturgy; the way it is shaped by emotion cultures; the consequences of deviating from permitted or expected scripts; and the dynamics of a socio-emotional economy, in which exchanges of emotion are affected by strategy, by power differentials and by gender, race and other characteristics.

References

- Acker, J. R. and Karp, D. 2006. *Wounds That Do Not Bind: Victim-Based Perspectives on the Death Penalty*. Durham, NC: Carolina Academic Press.
- Acorn, A. 2004. *Compulsory Compassion: A Critique of Restorative Justice*. Vancouver: University of British Columbia Press.
- Aguirre, A. et al. 1999. 'Sentencing Outcomes, Race, and Victim Impact Evidence in California: A Pre-and Post-Payne Comparison', *The Justice Professional* 11: 297–310.
- Armour, M. P. and Umbreit, M. S. 2006. 'Exploring "Closure" and the Ultimate Penal Sanction for Survivors of Homicide Victims', *Federal Sentencing Reporter* 19: 105–109.
- Arrigo, B. A. and Williams, C. R. 2003. 'Victim Vices, Victim Voices, and Impact Statements: On the Place of Emotion and the Role of Restorative Justice in Capital Sentencing', *Crime and Delinquency* 49: 603–626.
- Bandes, S. A. 1996. 'Empathy, Narrative, and Victim Impact Statements', *University of Chicago Law Review* 63: 361–413.
1999. 'Victim Standing', *Utah Law Review* 2: 331–348.
- 2000a. *The Passions of Law*. New York: New York University Press.
- 2000b. 'When Victims Seek Closure: Forgiveness, Vengeance and the Role of Government', *Fordham Urban Law Journal* 27: 1599–1607.

2006. 'Loyalty to One's Convictions: The Prosecutor and Tunnel Vision', *Howard Law Journal* 49: 475–495.
- 2009a. 'Repellent Crimes and Rational Deliberation: Emotion and the Death Penalty', *Vermont Law Review* 33: 489–519.
- 2009b. 'Victims, "Closure," and the Sociology of Emotion', *Law and Contemporary Problems* 72: 1–27.
- Bandes, S. A. and Salerno, J. 2014. 'Emotion, Proof and Prejudice: The Cognitive Science of Gruesome Photos and Victim Impact Statements', *Arizona State Law Journal* 46: 1003–1057.
- Barrett, L.F. and Bliss-Moreau, E. 2009. 'She's Emotional. He's Having a Bad Day: Attributional Explanations for Emotion Stereotypes', *Emotion* 9: 649–58.
- Bibas, S. and Bierschbach, R. A. 2004. 'Integrating Remorse and Apology into Criminal Procedure', *Yale Law Journal* 114: 85–149.
- Booth v. Maryland*, 482 U.S. 496 (1987).
- Bowers, W., Sandys, M. and Steiner, B. D. 1988. 'Foreclosed Impartiality in Capital Sentencing: Jurors' Predispositions, Guilt-Trial Experience, and Premature Decision Making', *Cornell Law Review* 83: 1476–1556.
- California v. Brown*, 479 U.S. 539 (1986).
- Clark, C. 1997. *Misery and Company: Sympathy in Everyday Life*. Chicago: University of Chicago Press.
- Davis, R. C. and Smith, B. E. 1994. 'Victim Impact Statements and Victim Satisfaction: An Unfulfilled Promise?', *Journal of Criminal Justice* 22: 1–12.
- Deise, J. and Paternoster, R. 2013. 'More than a "Quick Glimpse of the Life": The Relationship between Victim Impact Evidence and Death Sentencing', *Hastings Constitutional Law Quarterly* 40: 611–652.
- DeShaney v. Winnebago County*, 489 U.S. 189 (1989).
- Ekman, P. 1992. 'An Argument for Basic Emotions', *Cognition and Emotion* 6: 169–200.
- Ekman, P., Friesen, W. V. and Ancoli, S. 1980. 'Facial Signs of Emotional Experience', *Journal of Personality and Social Psychology* 39: 1125–1134.
- Erez, E. and Laster, K. 1999. 'Neutralizing Victim Reform: Legal Professionals' Perspective on Victims and Victim Impact Statements', *Crime and Delinquency* 45: 530–553.
- Greene, E., Koehring, H. and Quiat, M. 1998. 'Victim Impact Evidence in Capital Cases: Does the Victim's Character Matter?', *Journal of Applied Social Psychology* 28: 145–156.
- Hochschild, A. R. 1983. *The Managed Heart: Commercialization of Human Feeling*. Berkeley: University of California Press.
- Jacoby, S. 1983. *Wild Justice: The Evolution of Revenge*. New York: Harper & Row.
- Logan, W. A. 2008. 'Confronting Evil: Victims' Rights in an Age of Terror', *Georgetown Law Journal* 96: 721–777.
- Maroney, T. A. 2011. 'The Persistent Cultural Script of Judicial Dispassion', *California Law Review* 99: 629–682.
- McKinley Jr., J. C. 2014. 'At Sentencing Hearing, Killer Hears from Police, and His Victim's Mother', *New York Times* April 29: A21, www.nytimes.com/2014/04/30/nyregion/at-sentencing-hearing-in-manhattan-mother-lashes-out-at-daughters-killer.html, retrieved 3 October 2015.

- Miller, K. 2013. 'Purposing and Repurposing Harms: The Victim Impact Statement and Sexual Assault', *Qualitative Health Research* 23: 1445–1458.
- Mothers Against Drunk Drivers (MADD) Victim Services (2013). Victim Impact Statement Booklet, www.madd.org/victim-services/finding-support/victim-resources/victim-impact-statement.pdf, retrieved 7 June 2015.
- Paternoster, R. and Deise, J. 2011. 'A Heavy Thumb on the Scale: The Effect of Victim Impact Evidence on Capital Decision Making', *Criminology* 49: 129–161.
- Payne v. Tennessee, 501 U.S. 808 (1991).
- Richards, B. and Richards, D. 2015. 'To End the Anguish, Drop the Death Penalty', www.bostonglobe.com/metro/2015/04/16/end-anguish-drop-death-penalty/ocQLejp8H2vesDavItHIEN/story.html, retrieved 7 June 2015.
- Rose, M. R. and Diamond, S. S. 2009. 'Offstage Behavior: Real Jurors' Scrutiny of Non-testimonial Conduct', *DePaul Law Review* 58: 311–343.
- Rose, M. R., Nadler, J. and Clark, J. 2006. 'Appropriately Upset? Emotion Norms and Perceptions of Crime Victims', *Law and Human Behavior* 30: 203–219.
- Rossner, M. 2013. *Just Emotions: Rituals of Restorative Justice*. Oxford: Oxford University Press.
- Schuster, M. L. and Proppen, A. 2010. 'Degrees of Emotion: Judicial Responses to Victim Impact Statements', *Law Culture and the Humanities* 6: 75–104.
2011. *Victim Advocacy in the Courtroom: Persuasive Practices in Domestic Violence and Child Protection Cases*. Boston: Northeastern University Press.
- Seelye, K. Q. and Bidgood, J. 2015. 'Accounts of Heartbreak in Tsarnaev Trial as Victims of Boston Marathon Bombings Testify', www.nytimes.com/2015/03/06/us/-dzhokhar-tsarnaev-boston-marathon-bombing-trial-jeff-bauman.html?_r=0, retrieved 7 June 2015.
- Sundby, S. E. 2003. 'The Capital Jury and Empathy: The Problem of Worthy and Unworthy Victims', *Cornell Law Review* 88: 343–381.
- Thoits, P. A. 1989. 'The Sociology of Emotions', *Annual Review of Sociology* 15: 317–342.
- Victim Support Services (2015) <http://victimsupportservices.org/help-for-victims/victim-impact-statements/>, retrieved 3 October 2015.
- Zimring, F. 2003. *The Contradictions of American Capital Punishment*. New York: Oxford University Press.

13. Performing Anger to Signal Injustice

The Expression of Anger in Victim Impact Statements

Sarah Sorial

13.1 Introduction

Does the expression of anger have any role to play in victim impact statements and if so, what should that role be? Victim impact statements are written or oral statements made by the victim of a crime to jurors or judges as part of court proceedings.¹ The statements generally express the impact of the defendant's crime on the victim, or in capital cases, on the victim's family. The use of victim impact statements has become increasingly common in recent years, partly in response to concerns that victims have historically been neglected by the criminal justice system.² By being active participants in the legal process, victim impact statements allegedly bring victims closure and a sense that justice has been done.³

Many victims express anger at the injustice they have suffered, and some express a strong desire for vengeance or retribution. Although the expression of anger is understandable, given the gravity of the harms suffered by victims, judges and legal scholars have expressed some reticence or unease with expressions of anger in the context of a victim impact statement. Anger is considered a destructive emotion,

¹ Victim impact statements are legislated for in a number of different jurisdictions. In the United States, all fifty states and the federal government have enacted legislation recognizing a victim's right to provide a statement. See Wevodau et al. (2014: 46–47). Since 2005, they have been available to victims in The Netherlands, and are available in most Australian jurisdictions. For example, the New South Wales *Crimes (Sentencing Procedure) Act* 1999, Part 3, Division 2 ss 26–30A.

² For example, see Gewirtz (1996) for a defense of victim impact statements on the grounds that they allow otherwise silenced voices a hearing in the legal process.

³ See, for example, Chalmers, Duff and Leverick (2007).

which potentially undermines reasonable deliberation, impartiality and equality before the law.⁴

While taking into account these concerns, I suggest that there are good reasons for reconsidering the legitimacy of expressing anger in the context of a criminal trial and what the expression of anger is intended to achieve. In this chapter, I argue that expressions of anger in victim impact statements are invariably a kind of performance, intended to signal or communicate to specific audiences, such as judges and juries, that an injustice has occurred to a specific individual and that this injustice needs to be recognized and rectified. The issue then, is not so much one of whether victims should express anger, but *how* that anger should be expressed. To this end, I suggest that expressions of anger should be constrained by two important conditions: the expression of anger should be proportionate to the moral issue, and speakers should exercise restraint in the language they use to express their anger.

In [Section 13.2](#), I give an account of some of the features of anger and the different ways anger is expressed. In [Section 13.3](#), I examine some expressions of anger in victim impact statements and their effects on different intended audiences, such as judges and jurors. In [Section 13.4](#), I give an account of the relationship between law and the emotions in general, and law and the emotion of anger in particular, in order to explain the legal unease with expressions of anger in victim impact statements. This unease, I suggest, is partly attributable to some long-standing legal assumptions about the relationship between reason and the emotions, and the alleged destabilising effects of anger in particular. In [Section 13.5](#), I draw on some empirical evidence from social psychology that complicates this claim. The empirical evidence suggests that although intense anger may impair judgement, mild to moderate feelings of anger do not. In the case that anger might impair judgement, I suggest that the controlled nature of legal trials and courtrooms can act as safeguards to minimise these effects. In the final section, I suggest that victim impact statements need to be rethought as a kind of performance,

⁴ For example, many of the judges interviewed by Schuster and Proppen expressed unease with expressions of anger in victim impact statements (2010). Englebrecht and Chavez's (2014) empirical study found that prosecutors and victim advocates advised victims to remove expressions of anger from their statements. Legal scholars such as Susan Bandes have argued: 'victim impact statements are narratives that should be suppressed because they evoke emotions inappropriate in the context of criminal sentencing. Specifically, victim impact statements appeal to hatred, the desire for undifferentiated vengeance, and even bigotry' (Bandes 1996: 365–366).

where victims can perform or express anger in order to gain recognition for the harms or injustices they have suffered, and to have those harms redressed, subject to various qualifying conditions.

13.2 What Is Anger?

Anger seems to have the following features: first, it can feel a certain way. People feeling angry may have an increased heart rate; they may shake with anger or feel as if their 'blood is boiling' (Nussbaum 2001: 28–29). Second, these feelings typically arise in response to some perceived wrongdoing or injustice committed against an agent, although they need not (Aristotle 2003: 89).⁵ For example, irascible agents may often feel angry at trivial things that have nothing or little to do with any genuine wrongdoing. For these reasons, it is useful to distinguish between moral anger, which arises in response to legitimate wrongdoing or injustice, and general feelings of anger, which may have more to do with an agent's disposition. Moral anger is significant because it signals to the agent that some injustice has occurred or someone's rights have been violated (Prinz 2007: 38). Third, anger can be expressed in a variety of ways. People can express their feeling of anger by their facial expressions (by scowling, frowning, grimacing or gritting their teeth), by their tone of voice (such as shouting or raising their voice), their posture or by the words they use (Green 2007: 1). Fourth, and related to the third feature, anger can dispose people to certain actions. These actions are also expressions of anger, and can include such things as throwing or damaging things, or other acts of retaliatory violence (e.g., punching someone), although these actions do not always need to be expressions of anger (Green 2007: 1). Finally, there seems to be a strong connection between moral anger and the desire for retribution. For example, agents often express their anger in terms of a strong desire for vengeance (Aristotle 2003).

My focus in this chapter is on moral anger. Moral anger is complex because, on the one hand, it seems to do some important moral work in identifying injustice; on the other hand, the kinds of expressive acts to which it leads might not always be moral. There also seems to be an important difference between *feeling* anger and *expressing* that anger in

⁵ Aristotle thought that feelings of anger are not only justifiable in the face of injustice, but are also the mark of a virtuous person. A virtuous person feels the right amount of anger at the right time and for the right reasons (2003: 89).

various ways, including through violent retribution. It is the legitimacy of expressing anger, rather than feeling it, which concerns me here.

Expressions of anger, like self-expressions more generally, function to signal one's thoughts, affect or experience. It is how we show and share our point of view with others. As Green puts it, 'a state of mind is thus translucent in the sense that when I show it to you, you can become aware both of it and of what, if anything it represents' (Green 2007: 13). The expression or performance of how one feels not only enables others to be aware of what is going on, but also gives an indication of how the intended audience should appropriately respond (Green 2007: 14). In communicating or showing how an emotion or experience feels, those 'possessed of an appropriate level of empathy will then be able immediately, often even with no conscious deliberation, to identify with, succor, and so on, the creature who has shown them how they feel' (Green 2007: 14).

Expressions of anger in victim impact statements thus signal to relevant audiences how the victim has been affected by the injustice and how he or she feels about it. As such, they can be categorized as expressions of moral anger, rather than general anger. The anger is typically expressed through facial expressions or in words, rather than in violent action, although victims sometimes express a desire for violent forms of retribution. This expression not only seeks to communicate the harms that the victim has suffered but also functions to elicit empathy from others (such as judges and jurors), enabling them to properly gauge what is going on from the victim's perspective. More controversially, it also functions to prompt the audience to respond in the appropriate way by, for example, imposing an adequate penalty proportionate to the crime, in order to redress the injustice.

Given these features of anger, it is not surprising that victims of crime will feel angry in response to injustice, and will want to express that anger, given the opportunity. Whether they should *express* that anger in victim impact statements and the effects of this expression in the courtroom is a more complicated question, especially given the effects of this kind of testimony or narrative on various court actors, including judges and jurors.

13.3 Expressing Anger in Victim Impact Statements

Although different jurisdictions have different requirements or restrictions on the kind of information permitted, victim impact statements

typically contain information about the financial losses the person has suffered, changes to the victim's personal welfare or relationships and other psychological and physical harms. Victims express their feelings of grief, depression, anxiety, intense mourning at the loss of a loved one and their love for the victim. They also express anger, outrage and a desire for retribution. Some victims describe the defendant as an 'animal', a 'monster' or a 'creature'.⁶ Victims also express a strong desire for vengeance, either carried out by them or in the form of a severe sentence.⁷ Some victims express a strong desire for the death penalty, even in cases where the death penalty is not an option (see Bandes, this volume, for further discussion).⁸

In their analysis of judicial views of victim impact statements, Schuster and Proppen found that judges are generally uneasy about expressions of anger in the courtroom. Expressions of anger are thought to reflect immaturity, an inability to reach closure and a desire for personal revenge (Schuster and Proppen 2010: 83; for further discussion of which emotions are deemed appropriate and of the consequences of expressing inappropriate emotions, see Bandes and Weisman, both this volume). By contrast, expressions of grief are deemed acceptable as long as they are not excessive, as are expressions of compassion (although see Bandes, this volume, for discussion of cases, e.g., domestic violence, in which compassion is deemed inappropriate). The reasons for judicial unease about expressions of anger are that expressions of strong emotions, like anger, will undermine the impartiality of the legal process; will potentially undermine the principle of equality before the law; can sometimes undermine the authority of the judge and his or her control of the courtroom and can be directed at the judge or court officials if the victim is unsatisfied with the sentence (Schuster and Proppen 2010: 83).

⁶ For example, in transcripts of victim impact statements examined by Englebrecht and Chavez, one family member stated: '[The defendant] is not human. He may look like a man. He walks and talks. Outwardly he appears to be one of us, but the truth is he is not, for to be a human being means to possess a conscience, a soul, and his actions prove that he lacks this. Just look into his eyes as I have at every court appearance and I see no remorse, no regret, no sorrow, and no shame. [He] has proven himself a predator who preys on the weak and vulnerable'; and another statement said: '[The defendant] is an animal, a mad dog that should be put on a leash and caged for the rest of his natural life' (Englebrecht and Chavez 2014: 398).

⁷ For example, one victim concluded her statement by saying: 'If he goes for parole, I'm going to be waiting right outside the gates for him, that's all I got to say. And you're going to have another dead body up in here' (Englebrecht and Chavez 2014: 397).

⁸ Englebrecht and Chavez (2014).

Although all the judges who were interviewed acknowledged the importance of victim impact statements for victims, they also thought the statement should not affect the objectivity necessary for legal decision making. Legal decision making should be made rationally, and be based on facts and evidence, not emotion. Judges thus maintain that there is a strong demarcation between reason and emotion, and that the latter can corrupt the former in undesirable ways. Victim impact statements can potentially undermine reason by bringing strong emotions into the courtroom. For example, judges described often finding themselves in the awkward position of having to listen to victim impact statements that request particular sentencing conditions, while feeling constrained by standard sentencing requirements (Schuster and Propen 2010: 87).

Although judges do sometimes depart from standard sentencing guidelines in light of what victims say, they generally agree that to respond to a specific request for sentencing because of an impact statement can undermine some of the basic principles of the law, such as impartiality, objectivity and equality before the law (Schuster and Propen 2010: 88). Suppose, for example, that there were two identical murders by two different defendants. The family members of one victim give heart-wrenching accounts of their grief while the family of the other victim does not. Applying the law equitably and objectively means both the convicted parties should receive the same sentence, irrespective of the harms suffered and expressed by the victim's family.⁹ The introduction of strong emotion into the courtroom can distort people's perceptions of how the law ought to be applied and the principles underlying sentencing.¹⁰

In cases where the judge has not convinced a victim of why a particular sentence has been imposed, the anger can sometimes be misdirected at the judge.¹¹ Although judges do sometimes express their own

⁹ See the example given by one judge about the problem of imposing different sentences on the basis of how the crime has affected the victim's family (Schuster and Propen 2010: 87).

¹⁰ As one judge put it: 'If a judge asks, "what is it that you really want the court to do?" and the answer is "lock him up forever," or the answer is something completely contradictory of what the negotiation is, that can create a very awkward situation. Another judge said: 'if victims had their way, the defendant would be lowered one inch at a time into a wood chipper' (Schuster and Propen 2010: 92).

¹¹ According to one judge, 'The one thing that you see occasionally which I think is unfortunate is sort of the transfer of anger from being angry at the perpetrator to being angry at the court' (Schuster and Propen 2010: 93).

anger at the defendant, they also report feeling uncomfortable about expressing anger. Judges reported that anger makes them feel 'out of control' because it threatens both their judicial control of the courtroom and their neutral persona.¹²

There is a further concern among judges and legal scholars that victim impact statements might produce strong emotional reactions from jurors either because of the information conveyed or the affective demeanour of the victim giving the statement. Unlike judges, jurors do not have experience in criminal cases, or in sentencing, and so can be adversely affected by strong emotional expressions. According to Myers et al. (2002: 2394), victim impact testimony has the potential to influence jurors by two routes. One route is through the information about the harms suffered (referred to as *central persuasion*) and the other route is through the manner in which it is presented (referred to as *peripheral persuasion*) (see Bandes, this volume, for discussion of victim impact statements as sources of information).

The way the information is conveyed can be as powerful as the information itself, especially if there is variety of stimuli during a persuasion scenario. For example, jurors are more influenced by graphic material either in the form of photographic evidence or graphic descriptions. It might also be the case that jurors are influenced by the emotional presentation of the material. When presented with a distraught witness, jurors might become emotional themselves, and this could result in harsher penalties against the defendant (Myers et al. 2002: 2395). They might react to the emotions of others out of empathy or emotional contagion.¹³ The concern, then, is that strong emotions could be transferred from the witness to the jury and that strong emotions could influence subsequent judgements (Myers et al. 2002: 2396).

This concern is evidenced by several studies indicating that victim impact statements might, in fact, influence sentencing judgements. Wevodau et al. found that the presence of a victim impact statement resulted in a recommendation for lengthier sentences in noncapital cases (Wevodau 2014: 58). In capital cases, Luginbuhl and Burkhead found that mock jurors who read victim impact statements were more likely to choose the death penalty than those who did not (Luginbuhl and Burkhead 1995). Myers and Arbuthnot found that 67 per cent of

¹² See also for the effects of victim impact statements on jurors' assessment of evidence in Wevodau et al. (2014: 45–66).

¹³ See, for example, Sullins (1991).

mock jurors who witnessed victim impact evidence recommended that the defendant receive the death penalty compared to 30 per cent of those who did not receive victim impact evidence (Myers and Arbuthnot 1999). The question then is whether strong emotions like anger do impair judgement in this way and lead to harsher and more punitive punishment. The prevailing view in law is that it does, and so the expression of emotion more generally, and anger in particular, should not feature in victim impact statement. But this has not always been the case.

13.4 The Place of Anger in Law

Western law has had a complex relationship with the emotions more generally, and with anger in particular. The emotions, as Susan Bandes has argued, are either completely ignored in law, 'set aside', or only allowed in the context of a victim impact statement, and even then, it is acceptable only to express emotions like grief and compassion, but not anger or a desire for vengeance. As she puts it:

To label an influence 'emotional' is to say it is inappropriate – the very opposite of the reasoned discourse on which the legal system is premised. The traditional approach is to define those influences regarded as improper as irrational and thus emotional, and to treat influences that undeniably pervade the legal system as really not about emotion at all. (Bandes 2008–2009: 493)

Anger, being a strong emotion, is also considered an irrational emotion, because of the way it allegedly overwhelms and disrupts reason.¹⁴ Similarly, vengeance is thought to be disruptive, violent and unreasonable, and according to Robert Solomon, 'by its very nature, *opposed* to law and its constraints' (Solomon 1999: 129). However, the law's treatment of anger has historically been mixed.

For example, the ancient Athenians took seriously the problem of anger within the community. The primary function of the law was to address the source of this anger (the wrongdoing) in order to restore peace to the community. As Danielle Allen argues, 'anger was assumed to be not only at the root of the particular punishments but at the root

¹⁴ For example, in New Jersey, jurors are told: 'it is your duty to weigh the evidence calmly and without bias, passion, prejudice or sympathy, and to decide the issues upon the merits' ('New Jersey Criminal', 2007), while in North Dakota jurors are instructed: 'Your decision must not be influenced by sympathy or emotion' (State Bar Association of North Dakota 2008, cited in Salerno and Bottoms [2009: 276]).

of the whole structure of law itself' (Allen 1999: 194). Legal punishment existed because someone was angry at having suffered an injustice or wrongdoing and wanted the matter dealt with. It was the anger of the victim that made punishment necessary.

According to Allen, this was reflected in the structure of their penal practice. It was possible for any citizen to prosecute on behalf of another person or on behalf of the public, but in most cases, the prosecutor was either the victim or was personally involved in the case. The prosecutors usually justified their prosecutions on the grounds that they were seeking legitimate remedies for their anger. Anger was thus at the heart of the Athenian conceptions and experience of wrongdoing and punishment, to the extent that laws were established to help determine what levels of anger were appropriate for different kinds of wrongdoing (Allen 1999: 194).

There seemed to be a brief revival of this Aristotelian view of anger in early modern law. In his analysis of the changing conception of anger and law, Jeremy Horder notes that in the early modern period, there was an understanding that the passions had a positive ethical role to play and so should not be suppressed. With the guiding influence of reason, the passions could enable men to act virtuously. The early modern ethical theorists revived Aristotle's idea that cultivating the right sorts of passions in a man and performing the right sorts of actions constituted virtue (Horder 1992: 40). It was thought that there were circumstances where a person ought to be angry, just as there were circumstances where the person ought to feel afraid or compassionate. Acting virtuously consisted in showing a 'just and due proportion' not only in how one felt but also in how one acted, including in how one expressed one's feelings (Horder 1992: 41). This understanding of the relationship between virtue, anger and vengeance is also reflected in the requirements of 'honour-violence' and the law. Legally, expressing anger, even in the form of retaliatory violence, was deemed a rational response to various kinds of wrongdoing (Horder 1992: 42). This view was reflected in the early interpretations of the troubling defence of provocation.¹⁵

¹⁵ For example, the situations considered to be proper occasions for anger reflected the code of honour at the time. In *R v Mawgridge*, the court listed four categories of case that were 'by general consent' allowed as sufficient provocations. The first was the quarrel that escalated from words to physical assault. The second was a quarrel in which a friend of the person assaulted joined in and gave a deadly blow. The third was where someone took the part of a fellow citizen who was being 'injuriously treated' and the fourth was killing a man in the act of adultery with one's wife. See Horder (1992: 26).

In the eighteenth and nineteenth centuries, the law adopted a clearer distinction between reason and emotion, and the view of anger as reasonable 'outrage' was replaced with the idea that to be angry is to be 'out of control'. The prevailing view of the time, which persists to this day, was that when a person is calm, reason is exercising a 'dominion' or 'voice' from its 'seat or throne' in the soul. This calm was disrupted by the onset of anger (Horder 1992: 74). Anger was described as 'nothing but an ungoverned storm ... that men are subject to' (Horder 1992: 74). Great passions like anger were thought to deprive people of their reasoning abilities because the passions completely overwhelmed reason, rendering it useless in controlling a person's actions. Anger was thought to render a person 'beside himself' or to 'throw a man's mind off its balance' (Horder 1992: 75). An angry person is deemed to be 'out of control.' This view of anger has persisted in law, despite recent research contesting the view that emotions are inherently irrational and destabilizing forces. It is worth considering some of this recent empirical evidence, together with some recent developments, in the study of the emotions and law to give a more nuanced account of whether anger really does undermine reason and judgement in this way, and so undermines fair and equitable legal processes.

13.5 Anger, Judgement and the Empirical Evidence

The empirical evidence supports the theory that anger is intimately connected to injustice. Anger has been found to be associated with a sense or perception that the self (or someone close to the self) has been wronged, offended or injured in some way. The feeling of anger is also associated with a sense of certainty or confidence about what has happened and the belief that another person (not a random situation or the self) was responsible for the injury (Lerner and Tiedens 2009: 117). By contrast, negative events that are caused by forces outside an agent's control tend to evoke feelings of sadness rather than anger, and negative events caused by the self tend to evoke feelings of guilt or shame. When people feel uncertain or lack confidence about the cause of negative events, they are more likely to feel fear and anxiety than anger (Lerner and Tiedens 2009: 117).

The empirical evidence about the effects of anger on decision making and judgement is, however, inconclusive. Some of the evidence suggests that anger tends to impair rather than enable ethical decision making. Ethical situations are often complex, have no easy answers

and potentially serious consequences (Thiel et al. 2011: 380). They typically evoke strong emotional reactions for those involved. Gaudine and Thorne found that individuals experiencing positive affect were more likely to apply sophisticated cognitive moral structures to solve complex ethical situations (Gaudine and Thorne 2001). By contrast, experiencing a strong emotional state like anger increases the likelihood of making more punitive judgements of others involved in the ethical situation. The angry individual tends to have a more polarized view of other people, thinking that others are bad, while what they stand for is good (Thiel et al. 2011: 381). Bodenhausen et al. also found that people induced to feel anger engaged in more stereotyping than people induced to feel sadness. In persuasion paradigms, angry people paid little attention to the quality of the arguments made by the speaker, and instead, became convinced by superficial characteristics of the speaker (Bodenhausen et al. 1994).

As a consequence, angry individuals do not thoroughly evaluate ethical situations, make more punitive judgements and are more self-deceptive in making ethical choices (Bodenhausen et al. 1994). Part of the reason for this is that certainty can encourage optimism about future outcomes, which leads individuals to use less effort in evaluating the complex factors that promote ethical decision making. If people are certain of their chances or their abilities, they are less likely to systematically process information, given that this process allows for doubt and uncertainty (Bodenhausen et al. 1994: 385).

For example, when shown gruesome photographs of a murder scene, mock jurors gave more guilty verdicts and rated the prosecution's evidence as more sufficient than did jurors who decided on the same facts but viewed no photographs. These effects were statistically mediated by the mock jurors' reported feelings of anger towards the defendant. Viewing the photographs allegedly caused the jurors to become angry towards the defendant, and this anger led to the increase in guilty verdicts (Bright and Goodman-Delahunty 2006: 192). Anger can thus lead fact-finders to use evidence to make a decision on an emotional basis, give the evidence more weight than it should have and base their decisions on something other than the established propositions of the case (Bright and Goodman-Delahunty 2006: 200).

Moreover, because anger has a strong ability to capture attention, it affects other people's assessments of the angry individual. Although all emotional stimuli grab attention more so than neutral stimuli, anger has unique captivating properties for both the person experiencing it

and the person perceiving it. For example, Hansen and Hansen refer to the 'Anger Superiority Effect' – a tendency for people to pay particular attention to angry faces. Angry faces are detected easily and there are few mistakes made identifying an angry face as compared with neutral and happy facial expressions (Hansen and Hansen 1988: 919). Once people perceive angry targets, other, emotionally specific, inferences automatically follow: angry expressions are implicitly perceived as more threatening, competent, powerful and dominant, while sad expressers, by contrast, are perceived as being likable, submissive and in need of help (Lerner and Tiedens 2009: 116).

These findings give weight to the claim that anger may impair a juror's judgement and cause her to make a decision on the basis of an emotional response rather than a rational one. According to Bandes, there is evidence that jurors already enter the penalty phase with feelings of intense anger towards the defendant and intense sympathy towards the victim and the victim's family (Bandes 2008–2009: 501). This desire to punish may overwhelm jurors' capacity to evaluate the evidence rationally. Jurors thus need help in keeping their anger and grief from overwhelming their capacity to consider the defendant's arguments; victim impact statements may hinder this process by intensifying the pre-existing anger (Bandes 2008–2009: 501). In some cases, jurors may actively search for information that supports that conclusion or may distort events or facts in order to justify punishing the defendant (Lloyd-Bostock 1983). Individuals who become angry may seek an outlet for that anger in the form of a target they can punish (Myers and Greene 2004: 502).

There is, however, other empirical evidence, which suggests that although anger can bias thinking in various ways, it does not necessarily affect the quality or quantity of information processing. Moons and Mackie's study on whether anger hinders analytic processing found that angry people were sensitive to variations in the quality of arguments and message elaboration, irrespective of how the anger was induced. They also found that anger can 'override dispositional tendencies to process non-analytically and encourage more information processing' (Moons and Mackie 2007: 717).

The hasty decision making commonly associated with anger can be interpreted as an adaptive trigger for processing, designed to deal with environmental challenges. Anger, like other emotions, can trigger or motivate action in order to address an actual or potential threat; however, there is also value in accurately judging the source of the

threat and the most efficient way to deal with it when the circumstances allow. Thus an association between anger and analytic processing under certain circumstances may have adaptive benefits (Moons and Mackie 2007: 717).

Importantly, this study distinguished between different levels of anger arousal, including neutral anger and intense anger, which manifests itself as rage or fury. Although feelings of intense anger can inhibit analytic processing, mild and moderate anger do not. This finding is consistent with recent and compelling evidence that emotional judgments need not be irrational. It also has some important implications for our understanding of juror decision making in light of victim impact statements.

It is now well established that emotion is an essential source of information and influence on moral judgement. Some moral dilemmas engage emotional processing more than others, and these emotional engagements affect people's judgement (Greene et al. 2001: 2105). Emotions can highlight the salient features of moral dilemmas and enable us to understand other people, helping us to gauge their intentions and their credibility (Bandes 2009: 6). They also motivate us to care about the impact of our decision making and enable us to determine the most appropriate course of action in any given situation. Emotions may thus promote careful contemplation rather than interfere with it (Bandes 2009: 6).

Moreover, many emotions that the law has considered irrational and unstable – especially emotions like anger and vengeance – are, on closer inspection, quite rational. Anger often arises in response to a person's judgement that they or someone close to them has been wronged. Of course, a person's judgement and the resulting belief can be wrong, and the anger they feel and express may be completely irrational and misguided. It is also possible that expressions of anger can manifest as 'bursts' of uncontrollable emotion. However, the expression of anger in the context of victim impact statements is not one of these cases. Usually, the person expressing the anger is right in their judgement that they have been egregiously wronged, and so it is rational that they both feel and express that anger, provided it is proportionate to the moral harms. Given that the anger is expressed in a highly controlled courtroom environment, it is also unlikely that the anger will be uncontrollable or out of control.

Although the law perceives vengeance to be an irrational and out-of-control emotion, there is also a sense, as Robert Solomon has argued,

that it is a reflective and rational emotion, at least in the instrumental sense (Solomon 1999: 130). As Solomon puts it: 'vengeance in particular is both shaped and measured by way of reflection albeit (as in most emotions) a rather biased and narrow-minded reflection. Justice requires long term vision, but vengeance too is very different from simply "lashing out" and it is not at all blind' (Solomon 1999: 129). The function of law should not be to exclude these emotions, but to rationalize these powerful social passions (Solomon 1999: 131). I return to this point in the following section.

While jurors do make harsher judgements on the basis of emotional or graphic evidence, this might not be driven entirely by their emotional responses to gruesome evidence or angry impact statements. Jurors may well be processing the evidence analytically and rationally as well. As the evidence indicates, emotions and reasoning are not two distinct and independent spheres. According to the ruling in *Payne*, harsher sanctions that are arrived at this way are not prejudicial because they reflect a rational assessment of the harm rather than an irrational and arbitrary desire to punish (Myers and Greene 2004: 504). It may also be the case that mock jurors give harsher sentences in cases where they are shown gruesome or graphic photographs because they judge that the photographs have greater probative weight or value compared to a description of events. They may also feel angry, but that anger may be an accompanying feeling to the judgement about the value of the evidence. The anger may also be a perfectly rational response to their judgement that an injustice has occurred. Sentences may also be harsher because jurors might reason that a harsher sentence is warranted, given the level of harm that victims have suffered and communicated to the court.

Even if anger, especially intense anger, has the potential to impair judgement, as some of the evidence suggests, it is unlikely that anger is the only emotion being expressed in victim impact statements. Victims also express grief and sadness as a result of the crime so the picture about which emotions are being expressed and the effects that they have is much more complex. Emotions like sympathy and sadness tend to produce low arousal, insofar as agents feeling sympathy or sadness do not get as worked up or overwhelmed by the emotion, and so do not necessarily interfere with deliberate thought processes.

Moreover, in cases in which anger might impair judgement, there are various ways in which this can be mitigated in the trial process. A trial process encourages careful and thorough deliberation, both with others

and alone. Jurors need to discuss the case with one another, thereby requiring them to give justifications for their reasoning and decision making. Instructions and guidance by the presiding judge is intended to encourage contemplation and rational thought (Myers and Greene 2004: 503). This deliberative process may alleviate some of the deleterious effects of anger on judgement.

13.6 Performing Anger

The empirical evidence thus seems mixed about the effects of expressing anger in victim impact statements. There also seem to be at least two prevailing conceptions of anger in law: the view of anger as outrage in response to injustice and the view of anger as loss of control. I propose that there is a third way of thinking of expressions of moral anger in victim impact statements: this view is one of anger as a kind of performance. The idea that expression can have a performative dimension has a long history in the philosophy of language. As Austin has shown, many of the utterances that look like statements are not merely describing a state of affairs or expressing truth or falsity; rather these statements are also the performance of a certain action. For Austin, we *do* things with words. Our words can sometimes constitute various acts: in saying 'I promise' I actually make a promise; in saying 'I apologize' or 'thank you', I perform the acts apologising and thanking.

For Austin, an utterance has illocutionary force of a certain kind when it satisfies certain felicity conditions. These are typically set by conventions – written or unwritten – and typically require that the speaker is intending on doing something with her words (Langton 1993: 301). They further require that the speaker is the appropriate person to utter the words, and that the words are uttered by that speaker in the appropriate context. The illocutionary speech act is successful when the speaker is able to achieve audience 'uptake' or recognition of what she intends to do with her words.

I suggest that expressions of anger have an important performative dimension when they are expressed by the appropriate person (in this case, the victim), in the right context (such as in the context of a victim impact statement), for the following reasons: (1) the expression of anger can be used as a strategy to enable victims to secure recognition and redress for the harms they have suffered. In Austin's terms, the expression of anger can enable a victim to secure 'uptake'; (2) the expression

of anger can facilitate empathy; (3) it can elicit an appropriate response from audiences; (4) it can function, by giving expression to anger in the controlled environment of a trial process, it can function to restore community relations, which are often ruptured as a result of the crime (on the capacity of some emotions to restore community relations, see Helm, this volume). Thinking of victim impact statements in terms of their performative dimension (i.e., in terms of what they *do* rather than a mere description of how a victim is feeling) does, however, require a re-thinking of what the function of these statements is in the context of a criminal trial, and how that performance should be carried out.

13.6.1 Courtroom Performances

Despite stringent rules of evidence and procedure, the courtroom is a kind of theatre, where lawyers, defendants and victims perform various scripts and play out the guilt or innocence of the defendant for the adjudicators to assess. Each case develops a dynamic of its own, and the outcome is usually affected by various other factors that are not strictly evidence, including the quality of the lawyers' presentations, the appearance and demeanour of the defendant in the courtroom and the presence of the victims and their representatives (Levenson 2008: 574). Lawyers modify their own language, appearance and courtroom style in order to appeal to a jury. Defence lawyers also attempt to change how their clients look and behave, thereby encouraging them to perform their innocence before a jury.

Defendants do not appear in court in prison clothing or in shackles because their appearance as criminals can affect the jury's judgement and undermine the fairness of the trial (Levenson 2008: 584).¹⁶ Defendants are also encouraged to express the right kind of emotional responses, including remorse, regret and sadness for the crime they have committed. There is compelling evidence that these expressions or performances of these emotions influence how jurors assess the defendant. For example, Timothy McVeigh's failure to perform remorse or express any emotion was interpreted to mean he did not care. Although jurors claimed to not have discussed McVeigh's demeanour during deliberations, some noted individually 'they wanted to see more remorse' and that they were 'bothered that [McVeigh] was so stone-faced' (Levenson 2008: 596).

¹⁶ See also the ruling in *Estelle v Williams* 425 US 501 (1976).

13.6.2 *Performing Anger to Secure Recognition for Injustice*

If the courtroom is a kind of theatre where the relevant actors perform, then it follows that victims are also engaged in a kind of performance when they make impact statements. In the same way defendants perform or express remorse in order to demonstrate contrition, victims can also perform or express anger in order to signal to others that an injustice has occurred, to gain recognition for that injustice and to signal to others how they ought to respond. There are two kinds of recognition at stake here: institutional recognition by the court (represented by judicial recognition) and recognition of civil society or the community (represented by juror recognition).

The idea that anger has a performative dimension, which when used strategically, can enable people to gain recognition about an injustice is evident across a number of different spheres. For example, the performative dimension to anger is also apparent in peaceful protest marches, where groups of people use anger strategically to raise awareness about a particular injustice or to communicate something to their government. A failure to express anger might communicate that the people affected do not care enough or are indifferent to the issues. It is important to note that it need not matter whether the relevant parties actually *feel* anger, although it is likely that they would; the point is that in order to gain recognition for wrongdoing or injustice, it may be necessary to *perform* anger, by expressing it in various ways that are proportionate to the moral issue. In peaceful protest marches, the anger is usually expressed by shouting or holding up angry placards, which make various demands. Given that people pay more attention to angry faces, it might also be a good way of raising awareness about the issue by ensuring that others pay attention. Similarly, the expression of anger in the context of political debate can signal to audiences that an injustice has occurred, and that the angry person cares and is seeking some kind of redress.

Jesse Prinz has suggested that there are subtypes of anger depending on the eliciting conditions: anger is labelled indignation when elicited by injustice, and rage when elicited by a physical assault (Prinz 2007: 34). While I do not think these subtypes match up to eliciting conditions in quite this way (e.g., an irascible agent may fly into a rage over a trivial thing, irrespective of whether a physical assault has occurred, while an even-tempered person may feel rage over a matter of injustice), I suggest that these distinctions between the subtypes of anger

are useful for thinking about the ways anger can be expressed: anger can be expressed as indignation, or anger can be expressed as rage. Expressions of anger made in the context of peaceful protest marches or in the context of some political debate can be labelled indignation, while expression of anger like those made by victims examined can be labelled rage. I suggest that only performances of anger expressed as indignation are justifiable in the context of victim impact statements, for reasons I will explain in a later section.

The evidence about whether victim impact statements do in fact bring victims closure is inconclusive. It is also not clear that this alleged therapeutic function is appropriate for courtroom settings at all (see Bandes, this volume, 1996). But what might be appropriate for courtrooms and criminal trials in particular is a sense that justice has been served. One of the ways of achieving this is to ensure that victims gain adequate recognition for the injustices they have suffered. There is more compelling evidence that victim impact statements are in fact about seeking recognition, and that they are increasingly being treated as courtroom performances.

For example, Schuster and Proppen found that while many victims experienced a feeling of relief or satisfaction after making an impact statement, this was more attributable to the kind of judicial response they received (Schuster and Proppen 2010). Edna Erez found that ‘a major source of satisfaction for victims’ comes not from any direct effect on the sentence but comes ‘when judges pay attention to their input by citing victims’ own phrases from impact statements in judicial sentencing comments’ (Erez 1999: 553). Moreover, victim impact statements are also discussed with prosecutors and victims’ advocates, who make recommendations on the drafts they receive. Often these recommendations are to remove threatening language towards the defendant and angry remarks directed at the courts. Victims are also advised to respect the court and not to express vindictive, hateful or abusive language (Englebrecht and Chavez 2014: 400). This process suggests that victims, prosecutors and victim advocates perceive the impact statement as a performance, one that could reflect well or badly on them, and one that they want to get right.

13.6.3 Performing Anger to Foster Empathy and Elicit an Appropriate Response

Green has argued that one of the functions of showing how an emotion or experience feels is to put others in a position to empathize with the

person: 'we empathize with others when we imagine how they feel, but where the imagination in question must with reasonable accuracy capture that feeling ... once in your shoes I will then be more likely to come to your aid, assuage your pain, become any ally, and so forth' (Green 2007: 187). Expressing anger appropriately in the context of a victim impact statement can foster empathy for the victim, both in the form of recognition for the harms suffered and in enabling others to come to one's aid, perhaps by imposing an appropriate penalty on the offender. While many of the arguments against expressions of anger in victim impact statements have focused on the ways they affect jurors' judgments in capital cases, there is compelling evidence that the statements decreased victim blaming by jurors in sexual assault cases and led to lengthier sentencing recommendations by jurors (Wevodau 2014: 58). Given the prevalence of victim blaming in this area of law, and the low conviction rate in sexual assault trials, expressions of anger and other emotions in victim impact statements can foster greater empathy for the victim, and motivate others to respond appropriately by imposing sentences that are proportionate to the crime. The fostering of empathy can enable jurors to take seriously the harms suffered by sexual assault victims, and can contest some long-standing myths about sexual assault that have, historically, impeded justice for victims. Expressing or performing anger can thus enable better outcomes for victims.

13.6.4 Restoring Community Relations

Crime does not only affect the victims, but usually implicates the whole community in various ways. Expressions of anger and accompanying expressions of vengeance not only have an important communicative function for the victims involved, but they also have an important social dimension and public good. As Danielle Allen suggests, expressions of anger also function to 'spotlight' a problem of 'disordered relationality within the community, and points to the fact that wrongdoers, victims, and community all equally need to be cured' (Allen 1999: 197). Wrongdoing and punishment involves the community at large and produces a communal unease (or dis-ease, as Allen puts it). The expression of anger and the role of punishment also function to restore community relations and peace:

The form of 'retribution' employed in the Athenian context thus turns out not to be only about the 'just deserts' of a wrongdoer, not to be typically about the 'vengeance' of a victim but rather to be consistently about

recognising a disruption to communal peace and the need to restore that peace... All members of the community are in need of the restoration of peaceful and friendly relations within the community. All members of the community are mutually implicated in the problem of wrongdoing and its punishment. (Allen 1999: 197–198)

Re-thinking the function of anger in light of the role it played in ancient Athenian law reveals that wrongdoers are members of the community and their actions implicate that community; that punishment is a way of addressing disordered relationships within the community; that anger functions to ‘spotlight’ or ‘signal’ the need to take care of these disordered relationships; and that the law has a role in resolving that anger (partly by way of punishment) and restoring peace in the community (Allen 1999: 205).

With this analysis in mind, it is possible to interpret expressions of anger in victim impact statements as signalling or highlighting that an injustice has occurred, one that implicates the whole community and that others in the community ought to recognize it as such. The fact that jurors are so affected by victims’ emotions more generally suggests that the wrongdoing is not just about the victim, but the community as a whole. Serious crimes can make communities feel like their social cohesion has been destroyed; communities react with horror at the crime, feel sympathy and grief on behalf of the family and intense anger at the accused. Jurors, no doubt, bring some of these emotions into the courtroom (Bandes 2008–2009: 490). The penalties that jurors impose may also indicate their desire to restore harmony in a community where violence has occurred. The ancient Athenians often used extremely violent methods of punishment in order to restore peace to the community, and I am by no means suggesting that judges and jurors should do the same in addressing the problem of anger. I am, however, suggesting that expressions of anger may be justified in the context of a victim impact statement for the aforementioned reasons.

Expressions of anger should nevertheless be constrained by two important conditions: the expression of anger should be proportionate to the moral issue at hand, and speakers should exercise restraint in the language they use to express their anger. With respect to the first condition, speakers should clearly identify the moral issue (i.e., focus on the harms that have been done to them) and express their anger only in relation to that. It is inappropriate for speakers to misdirect their anger at court officials, the criminal process or the judge. While they may be angry at the process, the victim impact statement is not the place to

express this. Nor should speakers express anger to intimidate or disrespect court officials.

Second, speakers should use restraint in how they express their anger. If we think of anger as having subtypes, including subtypes of indignation and subtypes of rage, the expression of anger in victim impact statements should manifest as indignation, rather than rage. The problem with rage, as Nussbaum has argued, is that it can render its targets as seeming less than human, which in turn, can sanction various atrocities against them (Nussbaum 2001: 402–403). This can also address Susan Bandes's concern that the dehumanizing language used by victims to describe the defendant can have other, more pernicious effects. Although Bandes is critical of the mainstream view that the rule of law is somehow separate from emotions, and that emotion undermines legal reasoning, she nevertheless cautions against the expression of anger in victim impact statements because of the way this language might cause jurors to identify and empathise with the victim:

In evidentiary terms, victim impact statements are prejudicial and inflammatory. They overwhelm the jury with feelings of outrage toward the defendant and identification with the victim ... Victim impact statements diminish the jury's ability to process other relevant evidence, such as evidence in mitigation. (Bandes 2008–2009: 402)

They might also prevent jurors from perceiving the essential humanity of the defendant by interfering with, or blocking the juror's ability to empathise with, the defendant or conceive of him as human. Moderate expressions of anger, which manifest as indignation rather than rage, can potentially avoid this problematic outcome.

Victim impact statements have become an entrenched part of the criminal trial process in many jurisdictions. Given that moral anger arises in response to injustice, it is understandable that victims may want to express the anger they feel at the harms they have suffered. The issue, I have suggested, is not whether victims should express anger, but how that anger should be expressed. While there is a long-standing unease among legal scholars and practitioners that expressions of anger may erode long-standing legal principles, such as equality before the law, and may compromise legal deliberation, I have argued that expressions of anger may have an important and overlooked function in the context of victim impact statements. Expressions of anger are a kind of performance, which signal to others that an injustice has occurred to a

specific individual or group of people; they enable victims to gain institutional recognition for that injustice; they encourage the appropriate response on the part of others, in the form of imposing the appropriate penalty; and they can help restore relations in the community, which are often ruptured by crime. The point of a victim impact statement may not, then, be about closure for the victim; it may be more about recognition and redress for injustice.

References

- Allen, D. 1999. 'Democratic Dis-ease: Of Anger and the Troubling Nature of Punishment', in Bandes (ed.) *The Passions of Law*. New York: New York University Press.
- Aristotle. 2003. *The Nicomachean Ethics*. Translated by Thomson, J. A. K. London: Penguin.
- Bandes, S. 1996 'Empathy, Narrative, and Victim Impact Statements', *The University of Chicago Law Review* 63: 361–412.
- 2008–2009. 'Repellent Crimes and Rational Deliberation: Emotion and the Death Penalty', *Vermont Law Review* 33: 489–518.
2009. 'Victims, "Closure," and the Sociology of Emotion', *Law and Contemporary Problems* 72: 1–26.
- Bodenhausen, G. V., Sheppard, L. A. and Kramer, G. P. 1994. 'Negative Affect and Social Judgment: The Differential Impact of Anger and Sadness', *European Journal of Social Psychology Special Issue: Affect in Social Judgments and Cognition* 24: 45–62.
- Bright, D. A. and Goodman-Delahunty, J. 2006. 'Gruesome Evidence and Emotion: Anger, Blame, and Jury Decision-Making', *Law and Human Behavior* 30: 183–202.
- Chalmers, J., Duff, P. and Leverick, F. 2007. 'Victim Impact Statements: Can Work, Do Work (For Those Who Bother to Make Them)', *Criminal Law Review* May: 360–379.
- Englebrecht, C. M. and Chavez, J. M. 2014. 'Whose Statement Is It? An Examination of Victim Impact Statements Delivered in Court,' *Victims and Offenders: An International Journal of Evidence-based Research, Policy and Practice* 9: 386–412.
- Erez, E. 1999. 'Who's Afraid of the Big Bad Victim? Victim Impact Statements as Victim Empowerment and Enhancement of Justice,' *Criminal Law Review*: 545–556.
- Gaudine, A. and Thorne, L. 2001. 'Emotion and Ethical Decision-Making in Organisations', *Journal of Business Ethics* 31: 75–187.
- Gewirtz, P. 1996. 'Victims and Voyeurs: Two Narrative Problems at the Criminal Trial', in Brooks and Gewirtz (eds.) *Law's Stories: Narrative and Rhetoric in the Law*. New Haven, CT: Yale University Press.
- Green, M. 2007. *Self-Expression*. Oxford: Oxford University Press.

- Greene, J., Sommerville, B. R., Nystrom, L. E., Darley, J. M. and Cohen, J. D. 2009. 'An fMRI Investigation of Emotional Engagement in Moral Judgment', *Science* 293: 2105–2108.
- Hansen, C. H. and Hansen, R. D. 1988. 'Finding the Face in the Crowd: An Anger Superiority Effect', *Journal of Personality and Social Psychology* 54: 917–924.
- Horder, J. 1992. *Provocation and Responsibility*. Oxford: Clarendon Press.
- Langton, R. 1993. 'Speech Acts and Unspeakable Acts', *Philosophy and Public Affairs* 22: 293–330.
- Lerner, J. S. and Tiedens, L. Z. 2009. 'Portrait of the Angry Decision Maker', *Journal of Behavioral Decision Making* 19: 115–137.
- Levenson, L. L. 2008. 'Courtroom Demeanor: The Theatre of the Courtroom', *Minnesota Law Review* 92: 573–633.
- Lloyd-Bostock, S. 1983. 'Attributions of Cause and Responsibility as Social Phenomena', in Jaspers, Fincham and Hewstobe (eds.) *Attribution Theory and Research: Conceptual, Developmental, and Social Dimensions*. San Diego, CA: Academic Press, pp. 261–289.
- Luginbuhl, J. and Burkhead, M. 1995. 'Victim Impact Evidence in a Capital Trial: Encouraging Votes for Death', *American Journal of Criminal Justice* 20: 1–16.
- Luginbuhl, J. and Middendorg, K. 1988. 'Death Penalty Beliefs and Jurors' Responses to Aggravating and Mitigating Circumstances in Capital Trial', *Law and Human Behavior* 12: 263–281.
- Moons, W. and Mackie, D. 2007. 'Thinking Straight While Seeing Red: The Influence of Anger on Information Processing', *Personality and Social Psychology Bulletin* 33: 206–720.
- Myers, B. and Arbuthnot, J. 1999. 'The Effects of Victim Impact Information on the Verdicts and Sentence Recommendations of Mock Jurors', *Journal of Offender Rehabilitation* 29: 95–112.
- Myers, B. and Greene, E. 2004. 'The Prejudicial Nature of Victim Impact Statements: Implications for Capital Sentencing Policy', *Psychology, Public Policy and Law* 10: 492–515.
- Myers, B., Lynn, S. J. and Arbuthnot, J. 2002. 'Victim Impact Testimony and Juror Judgments: The Effects of Harm Information and Witness Demeanor', *Journal of Applied Social Psychology* 32: 2393–2412.
- Nussbaum, M. 2001. *Upheavals of Thought*. New York: Cambridge University Press.
- Prinz, J. 2007. *The Emotional Construction of Morals*. Oxford: Oxford University Press.
- Salerno, J. M. and Bottoms, B. L. 2009. 'Emotional Evidence and Jurors' Judgments: The Promise of Neuroscience for Informing Psychology and Law', *Behavioral Sciences & the Law* 27(2): 273–296.
- Schuster, M. L. and Proppen, A. 2010. 'Degrees of Emotion: Judicial Responses to Victim Impact Statements', *Law, Culture and the Humanities* 6: 75–104.
- Solomon, R. 1999. 'Justice v. Vengeance: On Law and the Satisfaction of Emotion' in Bandes. (ed.) *The Passions of Law*. New York: New York University Press, pp. 123–148.

- Sullins, E. S. 1991. 'Emotional Contagion Revisited: Effects of Social Comparison and Expressive Style on Mood Convergence', *Personality and Social Psychology Bulletin* 17: 166–174.
- Thiel, C., Connelly, S. and Griffith, J. A. 2011. 'The Influence of Anger on Ethical Decision Making: Comparison of a Primary and Secondary Appraisal', *Ethics and Behavior* 21: 380–403.
- Wevodau, A. L., Cramer, R. J., Clarke III, J. W. and Kehn, A. 2014. 'The Role of Emotion and Cognition in Juror Perceptions of Victim Impact Statements', *Social Justice Research* 27: 45–66.

Index

- Abrams, M. H., 85
Acorn, A., 283
action expressions. *See* [bodily expressions](#)
acts of speech, distinguished from speech acts, 35n19, *See also* [speech acts](#)
aesthetics, 2, 82
Affect Control Theory (ACT), 234
affect infusion model, 208
affect programs, 66, 66n25, 147, 158, 167
affective displays, of animals, 50n5
affective responses, discrete and automatic, 222
affective state, falsified portrayal of, 225
aggression, and perception of emotion, 209
Allen, D., 294, 295, 305
Alston, W. P., 97, 103
anger, 5, 228
 and analytic processing, 298–299
 courtroom performances of, 302
 and decision making, 297
 and ethical decision making, 296, 297
 expression of, in victim impact statements, 290–294
 and fear, 147
 features of, 289–290
 and gruesome evidence, 297
 and injustice, 296
 judicial unease about expression of, 291
 and law, 294–296
 and negotiation, 148
 and perception of emotion, 30
 performance of, 301–302
 role in fostering empathy and eliciting appropriate response, 304–305
 role in restoring community relations, 305–308
 role in securing recognition for injustice, 303–304
 subtypes of, 303
 victims' expressions of, 19
 of virtuous person, 289n5
Anger Superiority Effect, 298
angry faces, as enhancing cues of strength, 147
anthropology, of expression of emotions, 158
apology, 10, 11, 36, 36n21
 and remorse, 250–251
 ritual of, 250
appraisal respect, 102
appraisal theories, 8, 17, 160–161, 200, 201, 203
 of emotional communication, 225
appraisals, reverse engineering of, 203–205
approach-avoidance behaviour, 221
approach-related behaviours, 147
arational actions, 3, 74, 92, 98
Arbuthnot, J., 293
Aristotle, 157, 175
 on causality, 56
 view on anger, revival of, 295
art, nature and value of, 85
assertion, 11, 36
assertive family, 36n22
automatic expression/reaction, 162, 209
autonomic nervous system, 137, 158, 165, 166, 168
Aviezer, H., 210
avoidance-related behaviours, 147
Ayer, A. J., 49, 103

- Bandes, S., 294, 298, 307
 Bar-On, D., 12, 32, 33
 Barrett, L., 34
 Basic Emotion Theory (BET), 15–16, 157, 158, 159, 222
 and nature of emotion, 166–169
 prototype version of, 157–159
 revision to, 165–166
 basic emotions, 8, 9, 14–16, 26, 48, 67, 137
 insincere, 17–18, 226–229
 involuntary aspects of, 231–233
 voluntary aspects of, 229–231
 proponents of, 8
 sincere, 17–18, 226–229
 Baudelaire, C., 84
 behaviour, 53
 and emotional expression, 2, 3
 behavioural communication, 7
 behavioural components, of emotion, 29
 behavioural intention, 161–162
 behavioural marker, 33
 behavioural responses, 165
 belief–desire model of actions, 3, 98
 Betzler, M., 76
 Bibas, S., 275
 Biehl, M., 141
 Bierschbach, R., 275
 biological diversity, and emotional expression, 5
 biological emotions, 137
 and evolutionary theory, 138–139
 biologically innate emotions, 14
 bio-psych-social emotions, 137–138
 Blaison, C., 208
 blaming, 111
 blind vs. sighted individuals, emotions of, 143–145
 blindsight, Types I and II, 29
 blushing, 5, 270
 Bodenhausen, G. V., 297
 bodily changes, 12, 57–59, 66
 and emotion identification, 64
 bodily demeanours, 47, 48
 bodily expressions, 3, 4, 5, 30n10, 49n3
 bodily feelings, 120
 body posture, 173
 Bourke, J., 254
 Bruner, J. S., 210
 Burkhead, M., 293
 and remorse, 247
 and victim impact statements, 271–273, 293
 catharsis, 10, 19, 270, 271
 characteristic components, of emotion, 29, 33, 48, 51, 51n8, 52, 53, 53n11, 55, 56, 61n22, 69
 charged language, 37–41, 43
 children, capacity to regulate negative emotions, 151
 Clark, C., 267
 Clark, J., 282
 cognitivism, 58n18, 59–60, 59n20
 collective emotion
 approaches to, 116–119
 and shared emotion, 130–133
 collective emotional action, 14, 125–128
 collective emotional cognition, 14, 128–130
 collective emotional phenomena, 14, 119–122
 collective guilt, 117
 commitment, expression of. *See* expressions, of commitment
 common-sense view of emotions, 6, 47, 54–55, 57
 communicating emotions, 1, 5–9, 173, 178
 false emotion in high-stakes contexts, 233–237
 system of, 178
 communities of respect, 107
 complex process theories of emotions, 120
 compulsive self-reporters, of mental states, 33n16
 concealing true emotions, 229–230, 231
 confessional performances, in politics, 258
 confronting a defendant, 275
 connectionist models for visual cognition, 31n12
 conscious experience, 166
 constructivism, 210
 contempt, 229
 context
 dependent emotions, 229
 and facial expression, 17, 164, 205
 model of social signals, 211–213
 perceiver and, 206–209
 and perception of emotions, 199–202
 role in emotion perception, 209–211
 situational context, 205–206
 types, 205
 core affect (neurophysiological state), 34, 35
 ‘Correspondences’ (Baudelaire), 84

- corrugator supercilii, 230
- court, rationale for expression of emotions for, 276–278
- courtroom
 - demeanour, 302
 - emotion culture, 266–267
 - as a place for healing and closure, 269
 - victim emotion in, 268–269
- cross-cultural judgment/recognition of emotions, 141, 142, 174, 177, 179, 222–223
- cultural differences, in emotional
 - expressions, 5, 164, 166, 180, 186
 - facial behaviour, 140, 160
 - facial expressions, 15, 180–183
- cultural display rules, 149–151, 160, 166, 207–208
 - worldwide mapping of, 151–152
- cultural relativism, 176
- cultural script, of judicial dispassion, 273
- culturally learned expressions, 176–177, 210
- cultures, 148–149, 152–153
 - modelling facial expressions across, 187–190
- Damasio, A., 63, 64
- Darwall, S., 102, 108, 109, 111, 112
- Darwin, C., 3, 8, 15, 138, 139, 145, 146, 152, 158, 174, 175, 176, 177, 190, 199, 201, 209, 222, 226, 231, 232
- David, S., 206
- Davis, W., 5
- de Gelder, B., 206
- de Kock's case, 256–258
- death sentence. *See* capital cases
- deception, 221, 231, 235
 - voluntary and involuntary aspects, 224–225
- deceptive expressions of mind, 43
- deep acting (acting technique), 267
- defendant, rationale for expression of
 - emotions for, 274–276
- dehumanization, 260–261
- desired social behaviour, 147
- desires, 99
- Dewsbury, D. A., 138
- dimensional theories of emotion, 223
- discrete basic emotions, 4, 223, 231
- disgust, 228–229
 - facial expressions of, 175
 - types of, 210
- Display Rule Assessment Inventory (DRAI), 151–152
- display rules, 9, 150, 151
 - differences in, 152
- domestic violence cases, 282, 283
- Döring, S., 85, 86
- Dretske, F. I., 96
- Dual-Process Functional Theory (DPFT)
 - of remorse, 234
- Duchenne smile, 143, 213, 226, 231
- dynamic facial expressions, 16, 161, 187–190, 192, 213, 221
- easy facial expressions, 162
- Ekman, P., 4, 120, 121, 139, 140, 150, 158, 160, 164, 166, 174, 176, 177, 179, 180, 192, 199, 222, 223
- Eliot, T. S., 85
- Elkabetz, S., 211
- embodied cognition, 31
- emoticons, 186
- emotion recognition/identification, 16, 165
 - affecting factors, 203
 - cultural differences in, 16
 - gender differences in, 81n6, 208, 280, 281
 - two path-model of emotional recognition, 202–203
 - universal, 4, 14, 15, 16, 140, 141, 162, 164, 174, 176, 177, 179–180, 200, 222–224, 226
- emotional collective planning, 14
- emotional contagion, 209, 293
- emotional content accounts, 96–97
- emotional control
 - and self-monitoring, 233
 - and sincerity, 233
- emotional displays, 204. *See also* display rules
 - impact on observers, 147
 - rules for, 267
- emotional expressions
 - for conveying social information, 147
 - functional meaning of, 96
 - masking of, 17
 - nature of, 2–5
 - neutralization of, 17
 - normative significance, of, 9–10
 - perception of, 6, 12, 28–33, 43, 69–70, 140
 - and aggression, 209
 - and anger, 30
 - and context, 199–202
 - facial expressions, 17

- emotional expressions (*cont.*)
 - inferences from, 203–205
 - malleability limits, 209–211
 - part-whole, 51, 52, 55
 - and social status, 203
- phenomenological description of, 1
- and relationships among agents, 147–148
- simulation of, 17
- suppressing or mimicking a relevant, 10
- emotional intelligence (EI), 233
- emotional intentionality, 58, 58n18, 59n19
- emotional leakage, 8, 231, 232–233
 - identification of, 236, 237
- emotional liars, 225
- emotional phenomena, 115, 119n6, 122
 - collective, 14, 119–122
- emotional states, 67
- emotions
 - as behaviour motivators, 149
 - being expressive of, 3
 - characteristic components, 29
 - and ethical decision making, 296–297
 - experiential components of, 29
 - formal objects of, 57
 - and inner state, 63, 64
 - intentionality of, 7, 57, 82
 - judgements about others, 6
 - as judgments, 59
 - linguistic reports of, 103
 - literal perception of, 46
 - vs. mood, 42
 - non-inferential perceptual
 - knowledge of, 48
 - and outer state, 63, 64
 - particular objects of, 57
 - philosophical theories of, 57–60
 - physiological components of, 29, 33
 - as rapid information processing
 - systems, 138
 - social, 138
 - social functions of, 149
 - understanding of, 137–138
 - validly and reliably identified, 141
- empathy, 32, 32n15
- enabling knowledge, and showing
 - relevant to expression, 51
- endorsements of commitment. *See*
 - expressions, of commitment
- Erez, E., 304
- error-theory about common-sense
 - emotional categories, 34
- ethical decision making, and angry
 - individuals, 297
- ethical language, 2
- ethnicity, and emotional expression
 - evaluation, 280
- evaluation of emotional expression, and
 - race, 280, 281
- evaluative attitudes, 101
 - and dispositional emotion, 101n5
- evaluative stance, communication of
 - an, 8–9
- evincing and expressing emotions, 96–103
- evolution theory, 138–139, 175, *See also*
 - evolutionary theory
- evolutionarily adaptive emotions, 14
- evolutionary fitness, 96
- evolutionary theory, 8, 160–162, 166, *See also*
 - evolution theory
- expressions
 - and audiences, 271–272
 - causal sense of, 49n3
 - of commitment, 13, 14, 106
 - evaluation, of prototypical victim, 280
 - failures of, 56
 - genitive-or-generic account of, 25–28, 40
 - intensity and clarity of, 201
 - involuntary and spontaneous, 3, 7, 219–220, 224, 232
 - natural kinds of, 4, 33–35, 48, 49, 51, 96, 224
 - of psychological state, 27
 - of reactive attitudes, 13–14, 108–109, 108n8, 111, 112, 235
 - and representation, 28
 - semantic sense of, 49
 - as social signals, 199
- expressive actions, 12–14, 74, 92–93, 96, 97
 - and artefacts, 11
 - explanation of, 87–92
 - vs. expressive vehicles, 49
 - and justification of actions, 78–83
 - rationality of, 85
 - and reasons, 75–78
 - and symbolism, 83–87
- expressive artefacts, 40
- expressive autonomy, 39
- expressive behaviour, 47, 49, 53, 98, 121, 122–125, 137
 - and collective action, 127–128

- expressive figures, 41–43
- expressive theory of practical
 - rationality, 78
- expressive vehicles, 49
- expressiveness, 26, 27, 33n16
- expressivism, 39–40
- eyes
 - and facial expressions, 220
 - widening of, with fear or surprise, 146
- Facial Action Coding System (FACS),
 - 176, 223
- facial dominance, 221
- facial expressions, 1, 6, 8, 15, 47, 120,
 - 139–140, 190–192, 220
 - as adaptive features, 15
 - and affective response categories, 206
 - of basic emotions. *See* basic emotions
 - and body postures, 212
 - characterization of, 162
 - contortions, 48
 - controlling, for social interactions,
 - 229–230
 - cultural misinterpretation of, 174
 - culture-specific decoding of, 180–183
 - data-driven psychophysics approach to,
 - 183–187
 - and deception. *See* deception
 - and emotional states, 140
 - evolution of, 175–176
 - and false emotion communication,
 - 233–237
 - future research directions regarding,
 - 192–193
 - history of research into, 175–177
 - inaccurate reading of, 157
 - of infants, 222–223
 - innate, 162
 - and interpersonal effects, 146–148
 - and intrapersonal effects, 146
 - involuntary communication of
 - emotional information, 219
 - Japanese and Caucasian, 141
 - and job status, 207
 - obvious, 162
 - and other nonverbal channels, 205
 - perceptual expectation of, 183
 - production studies, 141–143
 - recognition network, 221
 - reliability of, 157
 - reliably recognized, 140
 - spontaneous
 - of blind athletes, 144
 - of blind individuals, 144
 - static, 16
 - status and role in interpersonal
 - communication, 220–221
 - studies of judgment of, 140–141
 - universal and learned expressions,
 - 222–224
- facial movements
 - onlooker's interpretation of, 162–165
 - production of, 159–162
- facial-expressive marker, 33
- facing surfaces, perceiving, 30, 51
- FACS. *See* Facial Action Coding System (FACS)
- false recognition, 163
- fear, 227–228
 - facial expressions of, 175
 - as a functional emotion, 227
 - perceived expressions of, 207
- feeling rules, 9, 248, 252, 267
- feeling theory, 57–59, 267
- feelings, 62, 63–64
 - hiding, 267
- felt evaluations, 99n4, 100–101
- Fernández-Dols, J., 34, 159
- figurative language, 43
- fittingness, of expressive vehicles, 77, 80,
 - 86, 89, 91
- fixation patterns, 181
- forensic psychology, 9–10
- Forgas, J., 208
- formal objects, of emotions, 57n18, 97
- Fridlund, A., 8, 161, 162
- Friesen, W., 4, 142, 150, 222, 223
- fusiform face area, 220
- Galati, D., 144
- Gaudine, A., 297
- gender differences, in emotion recognition.
 - See* emotion recognition / identification, gender differences in
- Generative Face Grammar (GFG), 188
- genitive-or-generic account of expression.
 - See* expressions, genitive-or-generic account of
- genuine and false expressions, variations
 - of, 232
- gestures, 48, 184
- Giere, R., 129

- Gilbert, M., 117
 Goffman, E., 79, 250
 Goldie, P., xi, 62, 63, 64, 88, 91, 99, 100, 121, 124, 132
 Green, M., 5, 11, 12, 26, 41, 50, 51, 52, 53, 54, 55, 56, 66, 67, 167, 290, 304
 Greenspan, P., 60
 Griffiths, P., 51, 66, 67
 group labels, for emotional expressions, 205, 208

 Hampton, J., 88, 89
 Hansen, C. H., 298
 Hansen, R. D., 298
 happiness, 226–227
 and smiles, 159
 hardwired emotion-specific brain circuits, 167
 Hareli, S., 17, 164, 206, 211
 healing, 270
 healing or catharsis, and expression of feelings in an open courtroom, 270
 Hess, U., 17, 164, 206, 208, 211
 Hochschild, A., 248, 267
 honour-violence, 295
 Horder, J., 295
 Horstmann, G., 159
 hostility, and perception of emotion, 209
 Hudson, A., 151
 Huebner, B., 117
 human and nonhuman animals, continuity between, 139
 Hume, D., 77n4
 Hunt, J., 258, 259
 Hursthouse, R., 3, 75, 76, 77, 85, 87, 89, 90, 92, 98
 Hursthouse's puzzle, 73–75
 hybrid theories of emotions, 60, 65, 67

 ideographic view of emotion, 34
 idiographic emotions, 34
 idiosyncratic expressions, 48
 characteristic components of, 51n8
 inanimate objects and processes, expressive properties of, 41–42
 inference, 12, 30–31, 52–54, 141, 203–205, 220
 based on behavioural evidence, 46
 epistemologist's notion of, 54
 psychologist's notion of, 54
 and situational information, 210
 unconscious, 31n12

 information processing, 208–209, 235
 information transmission and decoding, 177–179
 inhibition hypothesis, 231, 232, 236, 237
 initial impressions, durations of, 221
 insincere basic emotions. *See* basic emotions, insincere
 intentional action, 75–76, 90, 98n3
 intentional communication, 7
 intentional manipulation, 18, 230
 intentionality of emotion. *See* emotions, intentionality of
 interactive technologies, with social communication knowledge, 193
 interpersonal intentions, 204
 interpersonal perception, 234–235
 interpretations of facial expressions, 17, 165
 intimate killing, emotional impact of, 254–255
 involuntary and spontaneous expressions. *See* expressions, involuntary and spontaneous
 Izard, C., 4, 158, 200

 Jack, R., 9, 16, 180, 183, 184, 186, 188, 200
 Jacques, S., 151
 James, W., 57, 58, 59, 68, 69, 120
 Jamesian feeling theory, 59, 62, 63
 Japanese and Caucasian Facial Expressions of Emotion (JACFEE). *See* facial expressions, Japanese and Caucasian
 Jeshion, R., 38, 40
 job status, facial expressions and. *See* facial expressions, and job status
 Johnston, R. E., 160
 joint values, 95, 107–113
 judgements about others' emotional states. *See* emotions, judgements about others
 judgment formation, nature of, 60n21
 jurors
 evaluation of evidence, 298
 sensitivity of, 281, 297
 justification of emotion, 53, 59

 Kafetsios, K., 208
 Kayyal, M. H., 223
 Kivy, P., 41
 Klineberg, O., 160, 166
 Kraut, R. E., 160

- labelling emotions, reliability of, 223
 Landmann, H., 206
 Langfeld, H. S., 157
 Lazarus, R., 201
 learned expressions, 222–224
 legal theory, 9–10
 linguistic communication, 6
 linguistic reports, of emotion. *See* emotions, linguistic reports of
 literal perception of emotions, 51, 54, 55, 56, 60–69, *See* emotions, literal perception of
 Luginbuhl, J., 293
- Mackie, D., 298
 magical thinking, 88
 malice, 88
 masking of emotional expressions. *See* emotional expressions, masking of
 Matsumoto, D., 151, 152
 McDowell, J., 52
 mediation, victim–offender, 276, 283
 mental representations, 184
 mental states, and interactive technologies, 193
 metaphors, 12, 43
 Michael, J., 127, 128
 micro-expressions, 219, 232–233
 Mill, J. S., 84
 Miller, K., 274
 mindreading, 54n14
 minimal universality, 164
 moods vs. emotions. *See* emotions, vs. mood
 Moons, W., 298
 moral anger, 289
 and desire for retribution, 289
 moral communities, 18, 251–252, 260–261
 and remorse, 250–252
 occupations' influence on, 258–260
 regime change and remorse, 255–258
 undoing and rebuilding, in wartime, 253–255
 visibility of, 252–260
 moral dilemmas, and emotional processing, 299
 moral disgust, 210
 Mothers Against Drunk Driving (MADD), 271, 275
 motivational factors, 208
 Myers, B., 293
- Nadler, J., 282
 natural expressions, 82
 natural kinds of expressions. *See* expressions, natural kinds of
 natural selection theory, 176
 Nazis, regime change and remorse, 255–256
 negotiation, and anger. *See* anger, and negotiation
 neo-expressivist theory of avowals, 28n6
 neo-Jamesian feelings theory, 63, *See also* feelings theory
 neural adaptation, 221
 neutralization of emotional expressions. *See* emotional expressions, neutralization of
 non-capital cases, and victim impact statements, 273–274, 293, 305
 non-emotional states, 67
 non-inferential perceptual knowledge of emotions. *See* emotions, non-inferential perceptual knowledge of
 non-random responding, 165
 non-semantic expressions, 26
 norm knowledge, 207
 normative significance, of emotional expression. *See* emotional expressions, normative significance of
 Nussbaum, M., 307
- Oakley, J., 60
 objective correlation, 85
 occipital facial area, 220
 off-line simulation of mental life, 6
 orbicularis oculi, 230
 Ortega, J. E., 223
 outer state, and emotions. *See* emotions, and outer state
 overestimation, of intensity and clarity of expression. *See* expressions, intensity and clarity of
 overt affective experience, 224
 overt behaviour, 166
- pan-cultural behaviours, 48
 particular objects, of emotions, 57n18
 passions, ethical role, 295
 pattern-matching approach, 202, 203
 Payne, L., 258
 Peacocke, C., 47

- penal practice, and anger, 294–295
- perceived expression intensity, 141
- perceivers
 - goals, needs, and own emotional state in emotion identification, 208–209
 - as passive decoders, 202
- perception of emotional expressions. *See* emotional expressions, perception of
- and inference, 54n14, 55n15
- perceptual knowledge, 53
- performance or act of expressing, 49–50
- personal reactive attitudes, 108
- personality traits, and interactive technologies, 193
- pervasiveness of implicit rules for emotion display. *See* emotional display, rules for
- phenomenological description, of emotional expressions. *See* emotional expressions, phenomenological description of
- philosophical theories of emotion. *See* emotions, philosophical theories of
- physical disgust, 210
- physical/mental harm, and communication, 173
- physiognomy, 163–164
- physiological components, of emotion. *See* emotions, physiological components of
- Pictures of Facial Affect, 203
- pigmentocracies, 179
- plural agents, 107
- poetry and eloquence, distinction between, 84–85
- Porter, S., 225, 231, 233, 236
- post-Vietnam syndrome, 255
- predominant feeling rules, 267
- prejudice, 203
- primitive intelligibility, 99–100
- principle of serviceable associated habits, 15, 139, 146
- Prinz, J., 303
- Propen, A., 274, 282, 283, 291, 304
- proportionality norm, of intensity and duration of emotion, 282
- prototypical victim, expression evaluation of. *See* expressions, evaluation, of prototypical victim
- prototyping, 166–167
- psychological construction, 168
- psychological emotions, 137
- psychological state, expression of. *See* expressions, of psychological state
- psychopathic offenders, emotional actors, 220
- psychopathic traits, and unintended emotional leakage, 233
- psychopathy, 233
- psychophysiology, 183
- punitive judgments, of angry individuals, 297
- purposive expressions, 13
- race, and emotional expression evaluation. *See* evaluation of emotional expression, and race
- rage, 307
- rape cases, victim impact statement, 270, 274, 281, 282, 283
- rational explanation, 75
- rationality of emotion, 59, 75, 82
- Raz, J., 90
- reactive attitudes
 - expression of. *See* expressions, of reactive attitudes
 - vicarious, 108
- reason and emotion, 296
- recognition respect, 102
- Reisenzein, R., 159
- relationships among agents, and emotional expressions. *See* emotional expressions, and relationships among agents
- reliability of labelling emotions. *See* labelling emotions, reliability of
- reliably facial expressions. *See* facial expressions, reliability of
- remorse, 234–235, 254–255, 260–261
 - absence of, and death sentences, 247
 - and apology, 250–251, 275
 - false remorse, 235
 - genuine remorse, 235
 - mimicking of expressions of, 10
 - as a moral emotion, 234
 - and parole, 257
 - public expressions of, 253
 - and regime change, 255–258
 - regulative role of law, 18
- reports, expressions and, 13, 103–107
- representational acts and artefacts, 11

- resentment, 88
- restorative justice conferences, 268–269, 276
- retaliatory violence, 295
- reverse correlation model of perceptual expectations, 183, 184
- reverse engineering of appraisals. *See* appraisals, reverse engineering of
- Richard, M., 38, 39, 40
- Righart, R., 206
- rituals of inclusion, 250
- Roberts, R., 60
- Robinson, D. T., 234
- Robinson, J., 120, 121
- role expectations, 280
- Romanticism, 85
- Rose, M. R., 282
- Russell, J., 15, 16, 17, 34, 200, 223

- sadness, 227
- Salerno, J., 277
- Salmela, M., 118, 120, 121, 130, 131, 132
- Sartre, J-P., 88
- Scarantino, A., 166
- schadenfreude, 236
- Scherer, K. R., 205
- Schuster, M. L., 274, 282, 283, 291, 304
- Searle, J., 90
- secondary emotions, 207
- self-reactive attitudes, 108
- self-awareness, 2
- self-defeating, 88
- self-expression, 26, 27, 33n16, 290
- self-monitoring, and emotional control. *See* emotional control, and self-monitoring
- semantic sense, expressing in. *See* expression, semantic sense of
- Shannon, C. E., 201
- shared emotion, 118–119, 120, 130–133
- Shatan, C., 255
- showing, 27, 33n16, 42
 - expressive, inanimate phenomena as, 41
- showing-a, 27, 27n5, 32
- showing-how, 27
- showing-that, 27, 32
- Sias, J., 12
- signalling, 5, 32, 33n16, 36–37
- simulation of emotional expressions. *See* emotional expressions, simulation of
- Simulation Theory, 31
- sincere basic emotions. *See* basic emotions, sincere
- slurs/slurring, 11, 37–41, 43
 - expressivist accounts, 39, 40
 - non-offensive, 41n26
 - and words used slurringly, distinguish between, 38
- smile/smiling, 6, 160–161, 162, 166, 169, 199, 202, 208, 209, 213, 219, 226, 229, 230
 - and happiness, 159
- Smith, J., 30, 31
- Smith-Lovin, L., 234
- social class, 281
 - and emotional expression evaluation, 280
- social display rules, 150, 207
- social emotions. *See* emotions, social
- social functions, of emotions. *See* emotions, social functions of
- social group information, 205, 206
- social information, emotion expression for conveying. *See* emotional expression, for conveying social information
- social interactions, 140
 - emotional expressions in, 147
- social isolation, and communication, 173
- social judgement, spotlight effect in, 201
- social practices, 79, 251
- social referencing, 206
- Solomon, R., 294, 299, 300
- speech acts, 11, 28, 35–37, 43
 - sincerity-demanding instances, 37
- spite, 88, 89
- spotlight effect, in social judgment. *See* social judgement, spotlight effect in
- standardized universal facial expressions of emotion, 180
- states of mind, 25, 26, 36, 46, 48, 49, 50, 85, 137, 290
- static facial expressions. *See* facial expressions, static
- Stein, E., 1, 2
- stereotype expectation, 203, 207
- stimulus generation and task, 184–187
- Stout, R., 30, 51, 52, 55, 56, 67, 69
- Strawson, P., 108
- Studtmann, M., 159

- subjective perception of the stimulus, and appraisal, 203
- subjective/experiential marker, 33
- Sundby, S., 281
- suppressed desires, 91, 99
- suppressing or mimicking a relevant emotional expression. *See* emotional expressions, suppressing or mimicking a relevant
- surprise, 227
- survivors
 - guilt, 255
 - not supporting death sentence, 279
- symbolism, 86–87, 123–124
 - in art and poetry, 84–85
 - and expressive actions, 83–87
- sympathy exchanges, 267
- sympathy, and aid, 147

- Tagiuri, R., 210
- Tavuchis, N., 250, 251, 252
- Telles, E., 179
- ten Brinke, L., 225, 231
- theory of mind, 6
- Theory Theory, 31
- Thibault, P., 208
- Thorne, L., 297
- thoughts, and emotions, 132
- Tomkins, S., 4, 140, 158, 167
- transgressive act, transgressor's feelings about, 18
- trustworthiness identification, using facial features, 221
- two path-model of emotion recognition. *See* emotion recognition/identification, two path-model of emotional recognition

- unconscious inference, 31n12
- unfamiliar information, 178
- universal emotion recognition. *See* emotion recognition/identification, universal
- utterance, 3, 49
 - conventions, 301
- uttering, 39

- validly and reliably identified emotions. *See* emotions, validly and reliably identified
- vengeance, 294, 299–300
- vicarious reactive attitudes. *See* reactive attitudes, vicarious
- victim emotion
 - display rules, 278–284
 - explicit rules and expectations, 278–280
 - implicit rules, 280–284
- victim impact statements, 9, 19–20, 279, 287–289
 - examination of, 304
 - expressing anger in, 290–294
 - influence on jurors, 293
 - and maintaining neutral affect, 273–274
 - unsophisticated/colloquial, 282
- victim–offender mediation. *See* mediation, victim–offender
- victims
 - jurors' perceptions of, 281
 - need for neutral silence in courtroom, 274
 - rationale for expression of emotions for, 270–274
 - rights initiatives in favour of the death penalty, 268
 - role of expressions of emotion in legal and courtroom contexts, 18–19
 - worthy and unworthy, 264n2
- vocal expressions, 4, 47, 120
- voice tone, 47, 173

- Watson, G., 108
- Watson, J. B., 166
- Weaver, W., 201
- Weisman, R., 18
- Wenger, M. A., 166
- Wevodau, A. L., 293
- wish fulfilment, 124
- wish-frustration, 65n24
- wish-satisfaction, 65n24
- Wittgenstein, L., 49, 54
- Wollheim, R., 84, 86, 92

Titles Published in the Second Series (*continued from page ii*)

Emotions and Beliefs: How Feelings Influence Thoughts, edited by
Nico H. Frijda, Antony S. R. Manstead, and Sacha Bem

Identity and Emotion: Development through Self-Organization, edited by
Harke A. Bosma and E. Saskia Kunnen

Speaking from the Heart: Gender and the Social Meaning of Emotion, by
Stephanie A. Shields

The Hidden Genius of Emotion: Lifespan Transformations of Personality, by
Carol Magai and Jeannette Haviland-Jones

The Mind and Its Stories: Narrative Universals and Human Emotion, by
Patrick Colm Hogan

Feelings and Emotions: The Amsterdam Symposium, edited by
Antony S. R. Manstead, Nico H. Frijda, and Agneta H. Fischer

Collective Guilt: International Perspectives, edited by
Nyla R. Branscombe and Bertjan Doosje

The Social Life of Emotions, edited by
Larissa Z. Tiedens and Colin Wayne Leach

Emotions and Multilingualism, by
Aneta Pavlenko

Group Dynamics and Emotional Expression, edited by
Ursula Hess and Pierre Philippot

*Stigmatization, Tolerance and Repair: An Integrative Psychological Analysis of
Responses to Deviance*, by
Anton J. M. Dijkster and Willem Koomen

The Self and Its Emotions, by
Kristján Kristjánsson

*Face-to-Face Communication over the Internet: Emotions in a Web of Culture,
Language and Technology*, edited by
Arvid Kappas and Nicole C. Krämer

What Literature Teaches Us about Emotion, by
Patrick Colm Hogan

Emotion: A Biosocial Synthesis, by
Ross Buck

*From Self to Social Relationships: An Essentially Relational Perspective on
Social Motivation*, by
Martijn van Zomeren

*The Interpersonal Dynamics of Emotion: Towards an Integrative Theory of
Emotions as Social Information*, by
Gerban A. van Kleef

Emotional Mimicry in Social Context, edited by
Ursula Hess and Agneta Fischer

From Self to Social Relationships, by
Martijn van Zomeren

The Aesthetics of Emotion: Up the Down Staircase of the Mind-Body, by
Gerald C. Cupchik