

Scots Philosophical Association
University of St. Andrews

Explaining Action by Emotion

Author(s): Sabine A. Döring

Source: *The Philosophical Quarterly* (1950-), Vol. 53, No. 211 (Apr., 2003), pp. 214-230

Published by: Oxford University Press on behalf of the Scots Philosophical Association and the University of St. Andrews

Stable URL: <https://www.jstor.org/stable/3542865>

Accessed: 17-02-2020 21:30 UTC

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/3542865?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

Scots Philosophical Association, University of St. Andrews, Oxford University Press
are collaborating with JSTOR to digitize, preserve and extend access to *The Philosophical Quarterly* (1950-)

EXPLAINING ACTION BY EMOTION

BY SABINE A. DÖRING

I discuss two ways in which emotions explain actions: in the first, the explanation is expressive; in the second, the action is not only explained but also rationalized by the emotion's intentional content. The belief-desire model cannot satisfactorily account for either of these cases. My main purpose is to show that the emotions constitute an irreducible category in the explanation of action, to be understood by analogy with perception. Emotions are affective perceptions. Their affect gives them motivational force, and they can rationalize actions because, like perception, they have a representational intentional content. Because of this, an emotion can non-inferentially justify a belief which in its turn justifies or rationalizes an action; so emotions may constitute a source of moral knowledge.

I. INTRODUCTION

According to the currently predominant view, an intentional action is explained by an appropriate combination of belief and desire: agents perform a certain action because they desire something and believe the action to be a suitable means for attaining that thing. In combination with the belief, the desire gives them a reason for action, and when they act on this desire, the reason for their action is also its cause.¹ Opposing some recent claims, I shall argue that appeals to emotion in the explanation of action cannot satisfactorily be described in terms of the belief-desire model. As I shall show, the emotions constitute an irreducible category in the explanation of action, one that has to be understood by analogy with perception.

It is important to distinguish first two ways in which emotions can explain actions. In the first case, there is the expressive explanation of an action. In the second case, the action is not only explained but, by contrast with the expressive case, also rationalized by the emotion's intentional content. The belief-desire model fails to give a satisfactory account of either of these cases. In the expressive case, the belief-desire model rationalizes the action by attributing means-end reasoning to the agent where means-end

¹ See, e.g., Donald Davidson, 'Actions, Reasons, and Causes', repr. in his *Essays on Actions and Events* (Oxford: Clarendon Press, 1980), pp. 3-19.

reasoning, and thus rationalization, does not apply. In the case of rational explanation by the emotions, the belief–desire model either understates an emotion’s rationalizing function by reducing emotion to desire, or misrepresents this function in analysing it by analogy with belief. An emotion, rather, resembles a sense-perception in having an intentional content that is representational. As a consequence, an emotion can justify a belief. Like a perception, it can do so by its representational content non-inferentially justifying the content of that belief. If the emotion’s content provides an end for action, then the proposition, so justified, in its turn justifies or rationalizes an action. In that case the emotion functions as a non-inferential input to content-involving practical reasoning and to the explanation of rational action.

While I shall attribute an emotion’s rationalizing capacity to its representational content, the motivational force of an emotion has to be explained by reference to the emotion’s affect. Opposing the Humean theory of motivation, I shall claim that emotional motivation is independent of desire, that is, independent of whether the emotion provides an end for action. This is particularly clear in the case of expressive action.

II. GOLDIE’S DILEMMA

The explanation of expressive action is a touchstone for the belief–desire model, since this kind of action has the following distinctive characteristic: an action that is genuinely expressive of emotion is not performed by the agent as a means to some further end. Nor is the expressive action performed because it has some attractive feature, in the same way as throwing a ball can have the attractive feature of being fun. I shall call this characteristic ‘the absence of means–end reasoning’, understood as including the absence of explanation by attractive features.

A sophisticated attempt to apply the belief–desire model to expressive action has recently been made by Peter Goldie, who pursues Michael Smith’s account of the same subject further and thereby reformulates this account.² I shall argue that Goldie’s approach illustrates the shortcomings of any attempt to reconcile the belief–desire model with the absence of means–end reasoning characteristic of expressive action. In spite of Goldie’s efforts, explaining expressive action in terms of belief and desire remains distinctly unsatisfying.

² P. Goldie, *The Emotions: a Philosophical Exploration* (Oxford UP, 2000), ch. 5; see also his ‘Explaining Expression of Emotion’, *Mind*, 109 (2000), pp. 25–38. As the article is an extract from the book, I refer exclusively to the latter. See also M. Smith, ‘The Possibility of Action’, in J. Bransen (ed.), *Human Action, Deliberation and Causation* (Dordrecht: Kluwer, 1998), pp. 17–41.

Like Smith, Goldie rejects Rosalind Hursthouse's revisionary view that expressive actions belong to a class of so-called arational actions which are intentional, but which cannot be explained in the framework of Davidson's belief-desire account of action explanation.³ Though both authors agree that agents do not perform expressive actions as a means to some further end, they do not think that this is tantamount to accepting Hursthouse's view. Instead they claim that the belief-desire model can be modified in such a way that it provides a satisfactory account of expressive action. Hursthouse's example of Jane, who in a wave of hatred for Joan tears at Joan's photo with her nails, gouging holes in its eyes, is intended to show the validity of this claim.

Though Jane's expressive action cannot be explained as being intended as a means to some further end, a belief-desire explanation can be put forward for her action that identifies means and end. If Jane did what she did because she actually wanted to harm Joan and, being a believer in Voodoo, believed that she could achieve this by scratching out the eyes in a photograph of Joan, Jane's action would not be a genuine expression of an emotion. By contrast, Jane's action qualifies as expressive if she did what she did because she had the desire to scratch out the eyes in the photograph of Joan and believed that she could do this just by doing it. The necessity of attributing a belief in order to explain Jane's behaviour becomes clear when she acts on a false belief, e.g., scratching out the eyes in a photograph of June because she falsely believes that it is a photo of Joan. In this case Jane's action cannot be explained without reference to her false belief.

Both Smith and Goldie exhibit some unease with this explanation. Smith (p. 22) therefore suggests supplementing the belief-desire explanation by reference to an emotion, hatred, in this case, which does not supply Jane with a further reason but disposes her to act in a certain way. It is because Jane hates Joan that she is disposed to tear at Joan's photo with her nails, and to gouge holes in Joan's eyes. Against this explanatory strategy, Goldie objects that it is by no means clear why the emotion of hatred should dispose an agent to do such a bizarre thing as scratch out the eyes in a photo of a person she hates. As Goldie puts it, the desire to do this is not 'primitively intelligible'. He explains his notion of primitive intelligibility by saying (p. 43) that an emotional desire is 'primitively intelligible if it cannot be better explained by anything else other than the emotion of which it is a part'. Desiring to scratch out the eyes in a photo of a person one hates, Goldie says, is not the sort of desire that people in an occurrent state of hatred paradigmatically have. According to him, what has to be added to the belief-desire explanation of Jane's behaviour in order to render it

³ R. Hursthouse, 'Arational Actions', *Journal of Philosophy*, 88 (1991), pp. 57–68.

primitively intelligible, and thus by his lights make it an adequate explanation, is a further desire, namely, Jane's desire to scratch out Joan's eyes. Unlike Jane's desire to scratch out the eyes in a photo of Joan, her desire to scratch out the eyes of the real Joan is said to be primitively intelligible. It is said to be explicable only in virtue of Jane's hatred, of which it is supposedly a part. People in an occurrent state of hatred paradigmatically desire to scratch out the eyes of the hated person, on Goldie's account.

Why then does Jane's desire to scratch out the eyes of the real Joan move her to scratch out the eyes in a photograph of Joan? By stating that it does, Goldie implies that the belief-desire account of action explanation has to be extended so that it allows for the explanation of substitute actions, i.e., of actions springing from sublimated desires (*The Emotions*, pp. 129–36). According to this extended belief-desire account, Jane is doing what she is doing because she knows that she ought not to act on her desire to scratch out Joan's eyes, and she therefore 'symbolically satisfies', or rather sublimates, this desire by scratching out the eyes in a photograph of Joan. The cause of her action is further explained (p. 131) by a 'symbolic match or correspondence' between the object of her emotion and the object of her expressive action. The symbolic correspondence is that the photograph of Joan represents the real Joan to Jane. As Goldie emphasizes, Jane's action cannot be rationalized by reference to her desire to scratch out Joan's eyes. It would be a mistake to ascribe to Jane the belief that her desire concerning the real Joan would be fulfilled by doing anything to the photograph. On the contrary, if her action actually caused Joan to be blinded, Jane would most probably refrain from performing it. This is said to be so because Jane respects the constraints of modern civilization and is well aware of the fact that scratching out the eyes of someone does not quite fit these constraints. In the scenario outlined by Goldie, Jane's desire to scratch out Joan's eyes stems from an earlier, less civilized stage of human evolution: "Animal blood", to borrow Robert Musil's evocative phrase, has its residue in civilized adults, and this residue is revealed by the desires which are only "satisfied" in an etiolated, symbolic sense through the power of imagination' (*The Emotions*, p. 136). It is a principal concern of Goldie's to revise the belief-desire account of action explanation in a way that permits the description of us as agents whose desires are not perfectly rational, but stem at least partly from more primitive sources.

At this point Goldie's account is faced with a dilemma. Either the explanation of Jane's action is regarded as an instance of the belief-desire account or it is not. If it is, it must ascribe to Jane the desire to scratch out Joan's eyes in combination with two beliefs, first, that scratching out the eyes in a photo of Joan will at least allow her to vent her hatred towards the real

Joan on a substitute for its real object; secondly, the photo represents Joan. Under this description, however, Jane's action would not qualify as a genuine expression of an emotion but was rather intended as a means to the end of relief from emotional pressure. Expressive action is clearly over-intellectualized by such an explanation; the first belief needed is even some kind of meta-consideration. If the explanation proposed by Goldie is not regarded as an instance of the belief–desire model, then he has not made good his promise to treat expressive action within the belief–desire model.

To state that Goldie is facing a dilemma here is not to impugn all the elements of his account. The idea of mentally transferring an action-type from an object to its representation surely has some significant role to play in the psychology of some expressive actions. My point is that such a transfer across representational relations cannot, in the case of expressive actions, be captured within the belief–desire model. Rationality as captured in the belief–desire model cannot be reconciled with distinctively expressive action.

III. DO EMOTIONS IMPLY DESIRES?

In any case, not all examples of expressive action can be explained by Goldie's account. Not all expressive action involves symbolic representation. If a woman kicks the kitchen table after reading about the maintenance she has to pay for her ex-husband after their divorce, it would hardly make sense to claim that the kitchen table is a symbol for her ex-husband (or for anything else), instead of simply saying that the woman is venting her hatred or anger on the nearest thing to hand. But how can this less complex expressive action be analysed? All that Goldie says concerning this question is (p. 133) that if there is no symbolic correspondence between the object of emotion and the object of expressive action, 'the explanation is more likely to be ... due simply to being in the grip of an emotion'. This seems to acknowledge that whenever it is not possible to construe an expressive action as being the symbolic expression of a primitive desire, the belief–desire explanation of that action has to be supplemented just as Smith does supplement it, namely, by reference to an emotion. It seems that Goldie then has to deal with exactly the same objection as he initially raised against Smith. Why, if we are really trying to use the belief–desire model here, should someone's hatred or anger dispose him to do such a bizarre thing as kick the kitchen table?

Actually this same objection has to be raised equally against Goldie's claim that it is primitively intelligible that Jane's hatred for Joan should by itself immediately dispose her to scratch out Joan's eyes. The notion of

primitive intelligibility remains obscure. Goldie's view is that the desire to scratch out the eyes is a part of the hatred, and that the desire cannot be better explained by anything other than the hatred of which he says it is a part. But can we ever explain something merely by saying that it is a part of something else? We could hardly explain an event of a ball falling to the ground merely by saying that it is part of the larger event of two balls falling to the ground. Rather, when we do explain, we explain by specifying certain relations between the part to be explained and other parts of the whole. What Goldie does not tell us is the nature of the explanatory relations of the desire to scratch out the eyes to other events and states involved in hatred. We have yet to find the correct account of the relation that holds between an emotion and a particular action when the latter expresses the former.

It is not even clear what is meant by stating that the desire to scratch out the eyes of the hated person is a 'part' of hatred. What Goldie seems to be saying is that the primitive or 'brutish' desire is implied by the emotion-type of hatred. Because he is committed to the belief-desire account of action explanation, he seems to subscribe to the widely held view that emotions necessarily involve desires. Given the validity of the belief-desire account, an emotion can motivate the performance of an action only if it implies, or can even be reduced to, some desire. This is because the belief-desire account rests on the premise that agents' desires are the only states capable of motivating them to action; hence all explanation of action must start from an agent's desires. The motivational force of desire is attributed to a characteristic direction of fit called 'world-to-mind direction of fit', in contrast with the 'mind-to-world direction of fit' characteristic of belief.⁴ As opposed to beliefs, which are held to aim at truth, i.e., at fitting the world, it is claimed of desires that they aim at bringing about 'goals', and that is, conversely, at changing the world in such a way that it fits the desire. By virtue of their characteristic direction of fit or goal-directedness (rather than truth-directedness), desires are considered to be indispensable to the explanation of action, whereas beliefs are regarded as incapable of motivating: it is the desire, and not the belief, that provides an end for action. The desires agents have are claimed to dispose them to act in such a way that their goal is brought about, where the desired goal constitutes the end to which the action is intended as a means.

Prima facie it is false that emotions imply desires and thus necessarily provide ends for action. Emotions certainly do not imply that the world has to be changed in ways as specific as scratching out someone's eyes. An emotion need not provide an end for action at all. You may, for example, be

⁴ See M. Smith, 'The Humean Theory of Motivation', *Mind*, 96 (1987), pp. 36-61; and his *The Moral Problem* (Oxford: Blackwell, 1994), ch. 4.

proud of your achievement, sentimentally long for your former lover, or grieve over your mother's death, while at the same time lacking an end for action. In grieving over a beloved one's death, it is even impossible to change the world in such a way that it fits the emotion, since you cannot undo the person's death. For such reasons Goldie (p. 24) differentiates between typical desires and what he calls 'emotional desires', stressing that caution is needed 'when applying the idea of direction of fit to emotional desires'. As he admits, so-called emotional desires cannot easily be distinguished from other states by their world-to-mind direction of fit. Why then do 'emotional desires' qualify as desires? In fact Goldie gives no argument for subsuming so-called emotional desires within the general category of desire, but simply clings to the assumption that the belief-desire account of action explanation must be universally valid.

IV. AFFECTIVE PERCEPTION

So far it has been shown that even sophisticated attempts to force the explanation of expressive action into the belief-desire mould remain unsatisfying.⁵ In the next two sections my aim is to point out that there is no need whatsoever to do this, since these attempts all rest on a false presupposition. They each presume the so-called Humean theory of motivation, according to which all motivation is ultimately due to the desires an agent actually has. Opposing this theory, I shall be arguing that a distinctive motivational force is exhibited by the emotions. The hypothesis I shall defend is that an emotion's motivational force cannot be reduced to having a world-to-mind direction of fit, and therefore not to desire, but has to be understood in terms of what I shall call 'affective perception'.⁶

Nothing has been said so far about what an emotion is. The following characterization will prove useful: an emotion is an occurrent conscious state, with a certain affect, and with a certain kind of intentional content. This characterization is of course not meant to provide an exhaustive account of emotion. One could question whether all emotions are intentional.

⁵ Some would question whether expressions of emotions are actions. In my judgement, the general category of action is one with many subvarieties. I follow Brian O'Shaughnessy in holding that what makes something an action is that it is caused in a suitable way by one of the agent's tryings. Trying is present in the successful cases as well as in the failures. Anyone who tries to kick the table and fails because of some blockage of the neural signals would be surprised. See B. O'Shaughnessy, *The Will* (Cambridge UP, 1980), Vol. II, chs 9 and 11–15.

⁶ On the analogy between perception and the emotions see R. de Sousa, *The Rationality of Emotion* (MIT Press, 1987), pp. 149–56 and ch. 7; also L. Charland, 'Feeling and Representing: Computational Theory and the Modularity of Affect', *Synthese*, 105 (1996), pp. 273–301. Thanks are due to an anonymous referee for drawing my attention to Charland's paper.

This question concerns the so-called moods: apparently moods like anxiety or depression lack a target, since one can experience anxiety or depression without being anxious or depressed about anything in particular. For the purpose of this paper I shall not pursue this question further: I shall provisionally restrict the category of emotion to states fitting the given characterization, although I reject de Sousa's strategy (e.g., p. 7) of excluding moods from the class of emotions proper. This suffices to sketch an account of action explanation by the emotions which is distinct from any belief–desire explanation. In any case, 'the emotions' may not form a natural kind. From a psychological and neurobiological point of view, distinct classes may have to be distinguished among what are called 'emotions'.⁷ However, the elements of these classes share features which justify their provisional subsumption under one category in the philosophical theories of action and practical rationality.

Like a desire, an emotion is an intentional state: it is directed at a 'target' (see de Sousa, p. 116). Thus you hate your rivals, grieve over your mother's death, or are afraid of the aggressive-looking woman: your rivals, your mother's death and the aggressive-looking woman are the targets of your hatred, grief and fear respectively. In clear contrast with a desire, an emotion cannot be reduced to a merely functional state that disposes the subject towards action.⁸ This is because an emotion necessarily implies an evaluation of its target. Hating your rivals implies that you are seeing them as awful people; grieving over your mother's death implies that you are regarding her death as a sad event; being afraid of the aggressive-looking woman implies that you are thinking of her as dangerous. An emotion therefore has an intentional content that is evaluative.⁹

The thesis that emotions imply evaluations of their targets is often associated with the further view that each emotion-type has a so-called 'formal object' (corresponding to truth as the formal object of belief) which restricts and thereby determines the class of objects the particular type of emotion can be directed at. This formal-object view is required by the fact that the evaluation implied by an emotion is by no means arbitrary. In order for it to be a possible target of an emotion, the subject must see the object as

⁷ See, e.g., P. Griffiths, *What Emotions Really Are* (Chicago UP, 1997).

⁸ It is this dispositional conception of desires to which their interpretation in terms of their characteristic world-to-mind direction of fit amounts. See Smith, 'The Humean Theory of Motivation', p. 52.

⁹ Some claim that desires have an evaluative content too: see, e.g., G.E.M. Anscombe, *Intention* (Oxford: Blackwell, 1957), §§37ff.; de Sousa, 'The Good and the True', *Mind*, 83 (1974), pp. 534–51, and *The Rationality of Emotion*, pp. 20, 122–3; and T.M. Scanlon, *What We Owe to Each Other* (Cambridge: Belknap, 1998), pp. 38–41. I follow J. David Velleman in rejecting this claim. As Velleman argues, desires are directed at the *attainable*: 'The Guise of the Good', *Noûs*, 26 (1992), pp. 3–26, at p. 17.

having a certain property; otherwise the emotion would not be intelligible.¹⁰ Your fear of the aggressive-looking woman, for example, is only intelligible if there is some feature of her – looking aggressive, in this case – which explains why you see her as fearsome. If you were claiming to fear the woman while at the same time denying that there is anything fearsome about her, we would be conceptually excluded from understanding. Therefore it has been argued that the property which has to be ascribed to an emotion's target in order to render the emotion intelligible defines the emotion in question: this property is the emotion's formal object.

Though I agree with the formal-object view in principle, four qualifications need to be made. First, an emotion's intentional content more precisely is representational content.¹¹ Although representational content is also intentional, it differs from non-representational intentional content in being subject to a correctness condition. It is content that represents the world as being arranged in a certain way, and can thus be correct or incorrect. In experiencing fear of the aggressive-looking woman, it seems to you that the woman is in fact dangerous: the occurrent emotional state puts forward your fear's content as correct. By contrast, the content of an imagining is intentional, but people's imaginings do not present the world to them as being in a certain way. In imagining yourself to be Napoleon it does not seem to you, outright, that you really are Napoleon.¹²

Secondly, an emotion's formal object need not enter into the emotion's representational content. The ascription of an emotion's formal object to the target of that emotion may be merely implicit. Your fear of the aggressive-looking woman may simply represent her as dangerous. Fearsomeness then is implicitly ascribed to the target of your fear by this property's supervening on the woman's dangerousness; but this does not imply that fearsomeness features in the representational content.¹³

Thirdly, an emotion's representational content essentially differs from a belief's content. Emotions do not imply evaluative beliefs or judgements.¹⁴

¹⁰ See de Sousa, *The Rationality of Emotion*, pp. 122–3; W. Lyons, *Emotion* (Cambridge UP, 1980), pp. 99–104; A. Kenny, *Action, Emotion and Will* (London: Routledge & Kegan Paul, 1963), pp. 187–94.

¹¹ See C. Peacocke, *A Study of Concepts* (MIT Press, 1992), in particular ch. 3, and 'Scenarios, Contents and Perception', in T. Crane (ed.), *The Contents of Experience* (Cambridge UP, 1992), pp. 105–35.

¹² Robert C. Roberts is particularly clear on this. His view is that emotions are a certain kind of 'verisimilar concernful construals', and 'By "verisimilar" I mean to say that the construal has, for the construer, the *appearance* of truth, whether or not she would *affirm* the truth of the construal': 'What an Emotion Is: a Sketch', *Philosophical Review*, 97 (1988), pp. 183–209, at p. 191.

¹³ See also de Sousa, *The Rationality of Emotion*, pp. 122–3.

¹⁴ By contrast, Robert Solomon even claims that emotions can be reduced to evaluative judgements: *The Passions* (Indianapolis: Hackett, 1993), pp. 125–32.

Their content rather resembles the content of sense-perception in that both kinds of representational content need not be revised in the light of belief and better knowledge.¹⁵ The emotion of fear may represent a snake as dangerous. In other circumstances, fear may represent being high above the ground as dangerous. In both these cases, as Hume noted in the second, fear can persist, and still represent the snake or the height as dangerous, even when one judges, and knows, that the snake is harmless or that there is thick glass between oneself and the drop to the ground. Accordingly, it is not paradoxical, in the manner of Moore's paradox, but perfectly coherent to say 'I am afraid of the snake though (I know) it is not dangerous'. Moreover, our insight into our inclination to misunderstand other people's utterances as insulting need not lead to a change of our emotional reactions in future situations. If emotional evaluations were understood as evaluative beliefs, we would have to be described as holding inconsistent beliefs. But then we would be less rational than those who tend to misunderstand other people's utterances as insulting without realizing that in fact their own reactions are inappropriate. In order to avoid this unacceptable consequence, one has to distinguish between the content of emotion and the content of belief – just as the content of sense-perception has to be distinguished from the content of belief.

This is not to say that emotions can be reduced to perceptual evaluations. An evaluation can be present while the emotion is absent. As William P. Alston points out, 'two people can see a snake as equally dangerous ... and yet one is gripped with fear while the other is calm'.¹⁶ What distinguishes the emotion and makes it an affective perception is its feeling dimension, which is also called its 'affect'. This is my fourth and final qualification of the formal-object view: an emotion necessarily involves a certain affect.¹⁷

V. MOTIVATION WITHOUT DESIRE

Reference to an emotion's affective component is also required in the explanation of action by emotion. Explaining an action by emotion means specifying how the emotion's affect relates to its representational content in

¹⁵ E.g., for sense-perception, the Müller-Lyer illusion: see T. Crane, 'The Non-Conceptual Content of Experience', in his *The Contents of Experience* (Cambridge UP, 1992), pp. 150–1.

¹⁶ W.P. Alston, 'Emotion and Feeling', in P. Edwards (ed.), *Encyclopedia of Philosophy* (New York: Macmillan, 1967), Vol. II, pp. 479–86, at p. 485.

¹⁷ Opponents of this claim, who in my view tend to mistake the mere disposition towards an emotion for the emotion itself, are, e.g., E. Bedford, 'Emotions', in V.C. Chappell (ed.), *The Philosophy of Mind* (Englewood Cliffs: Prentice-Hall, 1962), pp. 110–26; G. Pitcher, 'Emotion', *Mind*, 74 (1965), pp. 326–46, at p. 338; Solomon, *The Passions*, pp. 96–102.

causing the action. It is the emotion's affect which gives it motivational force, rather than any desire being 'part' of it. Unlike a desire, an emotion's affect can still move its subject to act even if it is not necessary or actually impossible to change the world in such a way that it fits the emotion. In this case, which I claim to be the paradigm case of expressive action, the emotion's representational content fails to provide an end for action, which in turn means that the action cannot adequately be explained by means–end reasoning. In the example of Jane who, in a wave of hatred for Joan, scratches out the eyes in a photograph of Joan, the representational content of Jane's emotion is that Joan is an awful person. Seeing Joan as an awful person does not by itself provide an end for action. On the contrary, it may well be that there is nothing Jane can do in order to change the world in such a way that Joan no longer appears awful to her. Suppose Jane hates Joan because Jane's husband was unfaithful with and finally left his wife because of Joan. There is no way then in which Jane can undo what Joan did to her or what makes Joan an awful person in her eyes. Nevertheless hatred moves her towards action. Therefore she symbolically expresses her hatred for Joan by scratching out the eyes in a photo of Joan.

Imagine, by contrast, that Joan is afraid of the aggressive-looking Jane whom she unexpectedly meets in a dark street. The representational content of Joan's fear is that Jane is dangerous. As opposed to Jane's hatred for her rival, Joan's fear of the aggressive-looking Jane has a representational content that provides an end for action, namely, the end of avoiding danger. Any action that is thought by Joan to avoid the danger of the aggressive-looking Jane can be rationalized in this way, i.e., can be explained as a means to the end which is provided by the fear's representational content.

It follows from this contrast, between the action expressing Jane's hatred for her rival Joan and any action rationalized by Joan's fear of the aggressive-looking Jane, that whether an action springing from an emotion can adequately be explained in terms of means–end reasoning depends on that emotion's representational content. An emotion represents its target in a certain evaluative way, where the representation can, but need not, provide an end for action. As an emotion is still capable of motivating an action even if its representational content fails to provide an end for action, emotional motivation cannot be understood in terms of direction of fit. Instead the motivational force of emotion is to be explained in terms of the feeling-dimension of emotion: emotions are capable of motivating because their representational content is at the same time felt, i.e., because they are affective perceptions. Insisting, at least for those cases where an emotion's representational content does provide an end for action, that the emotion motivates because it has a world-to-mind direction of fit would clearly be

inconclusive. The motivational force of emotion would then be ascribed to a world-to-mind direction of fit only in some cases, whereas in other cases emotional motivation would have to be explained in terms of something else.

According to Charland, who also interprets emotions as affective perceptions, an emotion has motivational force because of being a 'two channel representational system' ('Feeling and Representing', p. 276). One channel is called 'exteroceptive' and is claimed to 'pick up and register affective information about the external world', while the other 'interoceptive' channel is said to be the means by which the subject 'is made internally aware of its internal physiological and somatic responses'. On Charland's account, the combined functioning of these two channels gives an emotion motivational force: 'You do not run from a predator unless you classify it as fearsome (exteroceptive affect), and you cannot run unless your physiological economy is properly mobilized for flight (interoceptive affect)'. Both the exteroceptive and the interoceptive channel are attributed to the feeling dimension of emotion. By introducing several empirical emotion theories and assessing their philosophical significance, Charland opposes a 'cognitivist' approach to emotion, arguing that 'affect is a distinct perceptual representation-governed system' (p. 273).

While I agree that emotions are perceptions that face two ways, I think that Charland's conceptualization of these two ways is inappropriate. What he calls the 'exteroceptive channel' classifies an emotion's target in terms of what I earlier called the emotion's 'formal object'. However, classifying a target as, say, fearsome already involves an 'interoception', i.e., a representation of something internal, for it implies that the subject attaches a certain import to avoiding harm (this is why, as Charland himself says, an emotional representation is *affective*). To classify something as fearsome means not just to distinguish it from other things in the external world but also to evaluate it, and that is to classify it with regard to the import it has for oneself.¹⁸ In so far as a target's import for the subject is represented in an emotional classification, the classification not only has an exteroceptive but at the same time an interoceptive dimension. Of course an interoception in this sense is not a perception of one's internal physiological states and processes. Emotions do involve certain physiological states and processes, but the import a target has for oneself in experiencing an emotion can by no means be reduced to a perception of these states and processes. I hold, *pace* Charland, that an emotional classification, or rather the representation of a target, is a cognition that has an interoceptive as well as an exteroceptive

¹⁸ To quote Roberts again, an emotion has a content that is 'concernful'. See his 'What an Emotion Is', p. 191.

dimension, where the interoceptive dimension must not be identified with a perception of one's internal physiological and somatic responses.

An emotional cognition is analogous to a sensual cognition in that its representational content cannot be separated from the emotion's conscious subjective character, i.e., from what it is like to experience, or rather to feel, the emotion. It is its affective character in which experiencing an emotion differs from experiencing a sense-perception: unlike a sense-perception, an emotion represents the target's import for the subject, and thus has an interoceptive dimension which makes it an affective perception and gives it motivational force.

Conceptualizing emotions as *cognitions* which face two ways is in fact compatible with the empirical theories of emotion quoted by Charland.¹⁹ It is true that all of these theories are directed against over-intellectualization of emotion. One of Charland's main sources is an influential paper in which the experimental psychologist Robert Zajonc claims to show that 'emotion', i.e., simple affective reactions, can exist before and independently of 'cognition'.²⁰ The crucial question is, however, what the term 'cognition' here refers to. As the title of his paper already indicates, Zajonc objects to the view that affective reactions depend on inferential judgement or belief. But he certainly does not deny that *some* information input is required in order to evoke a reaction at all.²¹ In my view, the information input required is processed into the emotion's representational content, which cannot be separated from the affective reaction itself and which makes this reaction a *distinct* cognition. As has been argued above, an emotion's representational content, like a sense-perception's representational content, has to be thoroughly distinguished from the content of judgement or belief, and need not, or rather *cannot*, be inferred from any other state's content.

The same defence of a cognitivist account of emotion can be put forward against Paul E. Griffiths: just like Charland, Griffiths regards the experiments reported by Zajonc as a refutation of this account, but confuses cognition with inferential judgement and belief.²² Griffiths' attack on cognitivist or 'propositional attitude' theories is directed against an account of emotion that (a) entirely relies on conceptual analysis, and (b) as a result of this inappropriate method 'makes some aspect of thought, usually a belief, central to the concept of emotion' (Griffiths, p. 21; quotation from Lyons,

¹⁹ I am indebted to Thomas Goschke for a number of very helpful discussions of the empirical literature on the emotions.

²⁰ R. Zajonc, 'Feeling and Thinking: Preferences Need No Inferences', in *American Psychologist*, 35 (1980), pp. 151–75.

²¹ This is also pointed out by Joseph LeDoux, *The Emotional Brain: the Mysterious Underpinnings of Emotional Life* (New York: Touchstone, 1998), p. 68.

²² See Griffiths, *What Emotions Really Are*, pp. 25–7 and 93–7.

Emotion, p. 33). Neither (a) nor (b) concerns 'cognitivism' as it is defended in this paper.

VI. ACTIONS RATIONALLY EXPLAINED BY EMOTION

In the previous section, explanations of action by emotion were divided into cases where the emotion's representational content provides an end for action, and cases where the action cannot adequately be rationalized in this way. It is important to emphasize that this division is a classification of explanations, not of the emotions themselves. One and the same emotion, such as fear, may on one occasion provide an end for action, whilst on another occasion it fails to do so. In contrast with Joan's fear in the example given above, one's fear may represent it as being dangerous to be so high up above the ground while one is travelling by plane and better off doing nothing at all. Furthermore, an emotion that provides an end for action may nevertheless be expressed. Joan's fear of the aggressive-looking Jane may even rationalize an action by providing an end for it, and explain an expressive action at the same time: while it may explain Joan's crossing the street as a means to the end of avoiding the danger of the aggressive-looking Jane, it may at the same time expressively explain Joan's clinging tightly to her bag. Although in the paradigm case of expressive action the emotion expressed does not provide an end for action, this is not to deny that emotions which aim at changing the world can be expressed.

Before I turn to the explanation of actions that are not only explained by an emotion but also rationalized by that emotion's representational content, some further qualification is needed concerning the distinction between this second class of explanations and the first class of explanations I have identified, the explanation of expressive actions. I shall call the second class 'rational', thereby implying that expressive actions are arational. The implication is justified by the fact that in the expressive case the agent ultimately lacks a reason for acting.

Expressive actions are rational in so far as the agent has to distinguish appropriate expressions from inappropriate ones. This is particularly important in cases where the action symbolizes 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 the representation includes the target's import for the subject. In the symbolic case, the rationality involved in expressive action consists of grasping the symbolic relations between emotional representations and their appropriate expressions. Jane's scratching out the eyes in a photograph of

Joan is in fact an example of a symbolic expression. Jane's action not only symbolizes the action of scratching out the eyes of the real Joan, but is also a symbol of her hatred's representational content, i.e., of Jane's seeing Joan as an awful person. This second symbolic function is indicated by the fact that acting in the way Jane does is an option reserved for women: only women are thought to express their hatred for another woman, typically a rival, by scratching out the other woman's eyes, whereas men are associated with actions like, e.g., breaking every bone in their rival's body. Scratching out the eyes of a hated person is part of a certain culturally established construction of femininity. Though the literal action of scratching out the eyes of a hated person is not a realistic option in our culture, it is this action that actually symbolizes the representational content of Jane's hatred. It follows that if Jane had scratched out the eyes of the real Joan, what she did would still have been a symbol of her hatred, and as such hardly the action of a brute beast. Contrary to Goldie's view, both scratching out the eyes in a photo of the hated person and scratching out the eyes of the real person one hates are culturally established symbols of hatred. The former action additionally symbolizes the latter, but it does not sublimate a desire stemming from 'animal blood'.

The most sophisticated way of symbolically expressing an emotion's representational content is achieved in art. Clearly the question whether expressing an emotion in a work of art is rational does not refer to whether the artist has a practical reason for creating that work of art. What is at stake here rather is the expression's appropriateness and quality as a symbol of the way the world appears to the subject in experiencing the emotion. As the exemplary case of artistic expression illustrates, the question whether an expression of an emotion is rational is a question of mind-to-world direction of fit, rather than of world-to-mind direction of fit. Having a representational content, an emotion essentially has a mind-to-world direction of fit. Only in those cases where an emotion at the same time has a world-to-mind direction of fit does it supply a reason for acting, and it does so only in a mediated way.²³ The claim I shall be finally defending for the second class of action explanations is that an emotion can give its subject a non-inferential way of justifying a practical proposition, which in its turn justifies or rationalizes an action. Since this capacity is not due to any belief involved in an emotion, nor can an emotion's motivating capacity be explained by the fact that some desire is 'part' of that emotion, the emotions constitute an irreducible category of practical reasoning. Accordingly, the belief-desire

²³ In claiming that an emotion can have *both* directions of fit at the same time I develop further an objection to Smith which was originally raised by David McNaughton. See Smith, 'The Humean Theory of Motivation', pp. 55–8.

model, though being of its essence designed to deal with rational explanations of action, also fails to give a satisfactory account for the explanation of actions that are rationalized by an emotion.

The fact that emotions have representational content implies that there is such a thing as taking them at face value – as taking the world to be as they represent it to be. This has two consequences for the rationalization of action by emotion: first, in being taken at face value, the representational content of an emotion can be an input to content-involving practical reasoning and to the explanation of action, without the need for any inference from the occurrence of the emotion. This is parallel to the fact that perceptual belief made rational by the representational content of the subject's experience is not a matter of an inference from the occurrence of the experience.²⁴ Secondly, the fact that emotions have representational content opens up the possibility that the occurrence of an emotion can, in suitable circumstances, entitle a thinker to judge, and possibly to know, its content simply by taking its representational content at face value. In the case of moral emotions, the possibility emerges that those emotions may give the thinker a non-inferential way of coming to know moral propositions.

Taking at face value the content of an emotion that has a world-to-mind direction of fit corresponds to a non-inferential justification of a practical belief's content. By 'practical', I mean that the belief's content implies the existence of a reason for acting, which in its turn justifies an action. For example, on being presented with the history of events leading up to the punishment of an individual, you may become indignant at its injustice. You may take the representational content of your indignation at face value, thereby coming to judge that the punishment was unjust. I assume that you would not have judged that content had you not experienced the indignation; otherwise your emotion would not be explanatorily significant. Provided further that moral propositions are practical, the proposition that the punishment was unjust implies the existence of a reason for acting.

The justificatory relation holds independently of whether you are actually entitled to take your emotion's content at face value: if you do rely on your emotion's content, you have a normative reason for acting. For the present I shall not address the question of what the conditions are under which a subject is entitled to take the representational content of an emotion at face value (it will hardly come as a surprise that, in my judgement, this question has to be answered by analogy to the entitlement in sense-perception). Instead, I shall consider finally the possibilities that would open up if there were a theory of entitlement in emotion. If those who take the

²⁴ See S.A. Döring and C. Peacocke, 'Handlungen, Gründe und Emotionen', in S.A. Döring and V. Mayer (eds), *Die Moralität der Gefühle* (Berlin: Akademie, 2002), pp. 81–103.

representational content of their indignation at an unjust punishment at face value were entitled to do so, they would have a normative reason for action that is justified by the fact that the punishment is unjust. The justificatory function of their emotion would be relatively superficial if they could have made the same moral judgement independently of their emotion, by using moral concepts they already accept. Although they did not in fact make the judgement of injustice prior to the emotion, they might well have so classified the punishment had the question arisen in thought. In a deeper kind of case, prior to experiencing the emotion, they would not have classified the punishment as unjust. Their indignation could still play a merely facilitative role in the justification of their judgements if, without their indignation, they would not in fact have judged that the punishment is unjust, but could in principle come to know that content by a route that does not involve the emotions. By contrast with this 'facilitative view', a stronger 'constitutive view' is that they could not have come to judge that the punishment is unjust except by endorsing the content of their indignation. The constitutive view would plausibly be entailed by some forms of virtue ethics. A virtue theory could, for instance, regard moral qualities as analogous to secondary qualities, and claim that the justice or injustice of something is its disposition to produce in thinkers the emotions that represent the thing as having that quality. In this case, moral values are accessible only by a route essentially involving the emotions, which therefore are indispensable to moral knowledge.

It has not yet been my aim to claim that the emotions indeed are a source of knowledge, but rather to provide a framework that allows proper articulation of this question. I hope that reflection on further examples of emotion-based judgement, and their role in our thought and action, will contribute, in the framework I have proposed, to an understanding of the relations between values and the emotions.²⁵

Universität Essen

²⁵ An earlier version of this paper was presented at the Philosophy Club, St Andrews University. I owe the idea put forward in §IV of this paper, that the intentional content of an emotion more specifically is representational content, to Christopher Peacocke, to whom I am deeply indebted (see 'Handlungen, Gründe und Emotionen'). I would like to thank the audience in St Andrews, Thomas Goschke, Thomas Grundmann, Fiona Macpherson, Christopher R. Taylor, Marcus Willaschek, and two anonymous referees, for valuable comments.