



This is a repository copy of *The problem of expressive action*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/169076/>

Version: Accepted Version

Article:

Bennett, C. orcid.org/0000-0001-8084-1210 (2021) The problem of expressive action. *Philosophy*, 96 (2). pp. 277-300. ISSN 0031-8191

<https://doi.org/10.1017/S0031819120000467>

© 2021 The Royal Institute of Philosophy. This is an author produced version of a paper accepted for publication in *Philosophy*. Article available under the terms of the CC-BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

The Problem of Expressive Action¹

ABSTRACT

Rational explanation of action out of emotion faces a number of challenges. The Wrong Explanation Challenge says that explaining action out of emotion by reference to a purpose rather than an emotion gets it wrong. The Redundancy Challenge says that if explanation of an action by reference to emotion is sufficient then rational explanation is redundant. And the No Further Justification Challenge says that there is no more to say, at the level of rational explanation, about why people act as they do out of a particular emotion. Furthermore, even if these challenges can be addressed, there is a Problem of Expressive Action, since many actions out of emotion seem unpromising candidates for being guided by normative practical reasons of the prudential, instrumental, deontic or consequentialist sort. In response, I argue that many actions out of emotions should be understood as expressive actions guided by the agent's conception of normative practical reasons: specifically, their conception of expressive reasons.

1. Problems in the Rational Explanation of Action out of Emotion

We are creatures of emotion. That is a commonplace. But the fact that emotion influences our conduct raises a range of important philosophical questions. Is action out of emotion (ever) rational action? Can action out of emotion (ever) be seen as an untrammelled expression of our rational agency, or only as its interruption or suspension? If the mark of rational action is responsiveness to reasons, are there reasons on which we act when we act out of emotion? Or is it rather the case that, if we do follow reason in acting out of emotion, we do so only indirectly, coincidentally, without any direct responsiveness to rational considerations?

¹ This paper was presented to audiences at the University of Valencia and the University of Pardubice, and I am grateful for the questions and comments that I received there. For discussion, I would particularly like to thank Maria Alvarez; Marina Barabas; Marta Cabrera; Josep Corbà; and Carlos Moya. I also received very helpful comments from two referees for this journal, for which I would like to express my gratitude.

As it will be understood in this paper, explanation of action in terms of rational agency portrays the action as being done on the basis of considerations that the agent took to count in favour of their action: that is by citing what the agent took to be reasons in the 'basic, normative sense,' as Scanlon puts it (Scanlon 1999; see also Alvarez 2010; Skorupski 2010). In this respect, rational explanation of action is no different from the rational explanation of belief. When we explain, for instance, why someone holds a belief by citing the evidence on which they formed that belief, we are citing the considerations that that person took to count in that belief's favour. The aim of rational explanation is to show how, as a rational agent, one came to believe or act as one did. In rational explanation, the belief or action is thus seen as the product of the agent's sensitivity to facts that stand – or were taken by the agent to stand – in a relation of rational support to it. Whereas rational explanation of belief sees us as oriented to what we can rationally take to be true, however, rational explanation of action sees us as oriented to what we can rationally take to be good or valuable. Clearly, we do not always act on optimal normative reasons; but explanation in terms of an agent's responsiveness to apparent normative reasons (that is, what the agent rationally takes to be the normative reasons that apply to them in the situation) is a key part of the way in which we make sense of the human world.

Not all human behaviour is susceptible to rational explanation, but I will take it that some of it is. How do things stand with respect to action out of emotion? On the one hand, there might seem to be a number of challenges to the idea of rational explanation of action out of emotion. First of all, there might seem a basic lack of fit between rational explanation of action and explanation in terms of the influence of emotion. If I kick my broken-down car in anger because I will now miss an important engagement, or I ruffle

my daughter's hair out of affection, it would be strange to say that I do it *for a reason*. Rather, it might seem, I do it *because of my emotion*. Normative reasons might seem to involve a deliberative motivational 'pull,' in contrast to emotion's spontaneous motivational 'push'. There are cases in which I *might* do it for some further goal – for instance, I might kick the car in order to release my tension, or I might ruffle my child's hair in order to make her understand that she is loved – but these are not the most common or most fundamental cases. We can call this the Wrong Explanation Challenge: explaining action out of emotion by reference to a purpose rather than an emotion gets it wrong.

Furthermore, reference to the emotion seems to provide a *sufficient* explanation of such common actions. As Kovach and de Lancey put it:

'For certain behaviors, including some emotional actions, given some (often minimal) background knowledge about a person's situation and the knowledge that the person is in the grip of an [emotion], nothing further is needed to explain why a person engages in a behavior.' (Kovach and de Lancey 2005: p. 114)

If rational action is explained by our responsiveness to reasons then it looks as though rational explanation is redundant given that the presence of emotion provides a sufficient explanation. Call this the Redundancy Challenge: if explanation of an action by reference to emotion is sufficient then rational explanation is redundant.

And finally, the explanation of action as done out of emotion might be thought to lack the characteristic structure of rational justification. When we ask why someone has

turned on the oven, for instance, and answer that they want to bake a cake, their goal explains the action; but we can also seek a *deeper* explanation by repeating the question and ask why they want to bake a cake. If the answer is that they are having guests and want to show hospitality then we can again ask why they want to do that. The chain of justifications will run out at some point, of course, but rational explanation of action has an in-built potential for depth, whereby immediate goals are explained by further goals that ground them. However, the explanation of action by reference to emotion seems to preclude such depth. We might be able to explain what counts in favour of a given emotional attitude, given a particular situation. In doing so we might explain why the attitude was fitting to the situation, given the formal object of the emotion.² But can we say anything about why, given that they were experiencing that emotion, an agent *acted* as they did? When we ask why a person in the grip of anger kicks their car, it might seem that there is nothing to say other than ‘that is the kind of thing that people do when they are angry’. The link between emotion and action might thus be thought simply to be, as Goldie has argued, ‘primitively intelligible’ (Goldie 2000). Insofar as there is something more to say, this would be at the level of non-rational rather than rational explanation. Call this the No Further Justification Challenge: that there is no more to say, at the level of rational explanation, about why people act as they do out of a particular emotion.

On the other hand, however, some important considerations suggest that explanation of action out of emotion is more continuous with rational explanation than these challenges would allow. For one thing, emotion is pervasive, and if action out of emotion

² Though we might need more than mere fittingness to capture the idea that there is something that counts in favor of the emotion. See (D’Arms and Jacobson 2000).

were always explained non-rationally then we might think there would be few actions for rational explanation to explain. Yet rational explanation, in some form or another, seems a central part of our understanding of one another. Furthermore, while the challenges suggest a strict dichotomy between rational action and action out of emotion, it might appear, pre-theoretically at least, quite hard to tell, either first- or third-personally, whether someone's motivation was emotional or rational. This might cast doubt on the idea of a strict dichotomy, or at least suggest that it would need to be independently motivated. And finally, we hold one another accountable for action out of emotion, blamed when it is inappropriate and praised when it is appropriate. While it is not necessarily true to say that an action that cannot be explained rationally is one for which we cannot be appraised, if we think of accountability as a process of holding a person to a rational standard that they could and should have met, there will be a strong connection between accountability and a form of explanation that assumes our sensitivity to rational considerations (Smith 2005; Bennett 2012).

In response to this contradiction, it might be argued that a range of things can be meant by 'action out of emotion,' and that each of the points noted above might be true of some types of such action and not true of others. We recognize interruptions to our rational agency and acknowledge that physical and psychological forces can leave our agency overborne. For instance, Rosalind Hursthouse notes that the influence of emotion on behaviour can range from reactions that are uncontrollable and involuntary; through reactions that we can prevent, either temporarily or permanently, through an act of will; to patterns of action and reaction that are inculcated through education and which, though hard to change as a result of reasoning, may be judged appropriate or inappropriate; to acts of emotion where the emotion is deliberately cultivated; and so

on (Hursthouse 1991). However, while this point is well taken, it would not explain how the three challenges I have just outlined could be met. What we need is an explanation of how it is possible that some action out of emotion can be guided by normative reasons to which the agent is responsive in so acting.

A better response is to argue that the emotions play a role in *reinforcing or disclosing* normative reasons (Greenspan 1988; Greenspan 2000; Helm 2001; Carman 2018a; Carman 2018b). For instance, the emotion of fear, it might be argued, can highlight prudential reasons in situations of danger; the emotion of compassion can highlight reasons based in the needs of others. However, even if this is correct, the range of emotional actions that could be shown to be rational by such a strategy looks to be limited. If the assumption is that the normative reasons reinforced or disclosed by the emotions are those that are grounded in widely recognized canons of prudential or instrumental rationality, or in a conception of deontic or consequentialist normative reasons, we will continue to struggle to make sense of the rationality of any actions for which none of these widely recognized rational grounds obtain. Yet there are many such emotional actions. Some examples are given by Hursthouse: ruffling a child's hair out of affection; jealously stabbing out the eyes of a picture of a rival; etc (Hursthouse 1991). Another example might be the anger of oppressed peoples, which, Amia Srinivasan has argued, can prompt action that conflicts with what such agents have prudential reason to do in their situation (Srinivasan 2018). Sensing, perhaps, that not all such actions can merely be dismissed as irrational, but being inarticulate about what could possibly count in their favour, we tend to talk of them as *expressive*. Some theorists, furthermore, have defended expressive action as a valuable form of response (Anderson 1995; Anderson and Pildes 2000). However, for many, it remains unclear how the fact that an

action is expressive can count rationally in its favour in the absence of rational grounding in those widely recognized rational grounds that we mentioned earlier.³ This suggests that the task of showing the rationality of action out of emotion faces a further challenge: the Problem of Expressive Action. This is the problem of understanding how action out of emotion can be rational even though its rationality cannot always be grounded in prudence, or instrumental rationality, or the deontic or consequentialist considerations highlighted by standard versions of normative theory.

I will argue that the debate on the range and rationality of action out of emotion has thus far missed out an important way in which much action out of emotion can be seen as guided by considerations that the agent takes to favour so acting. To explain this, I will argue for the rationality of what I will call *expressive action*. Expressive action conforms to canons of rational explanation because it is guided by the agent's conception of normative reasons: that is, of considerations that the agent takes to count in favour of so acting. Expressive action, however, is not done *for a reason* in the sense of aiming at some further end; it is rather taken to be its own end. However, we should not conclude that expressive action therefore has no rational basis. Expressive action aims, not to change the situation, but rather to *do justice* to it by virtue of the *expressive power* of the selected action (Bennett 2016). Expressive action is thus guided by a conception of expressive reasons: reasons are reasons to do justice to a situation by acting in a way that is expressively powerful in relation to the salient features of that situation. Expressive reasons would be distinct from instrumental, prudential, deontic

³ See e.g. Jason Brennan and Peter Jaworski's arguments against what they call 'semiotic objections' to certain market activities in (Brennan and Jaworski 2015).

or consequentialist reasons, and would be a *sui generis* type of normative reason that the agent takes to favour acting as they do.

If this account is valid, we would have a solution of the Problem of Expressive Action that would show how expressive actions out of emotion can be candidates for rational explanation. It would show that rational explanation does not give the Wrong Explanation. It would also show that the idea that there is No Further Justification for action out of emotion is false: expressive action would be seen as guided by an agent's conception of normative reasons to perform a certain action as an expressively powerful response that does justice to the significance of the agent's situation. My claim is thus that action out of emotion that might otherwise be labelled as *merely* expressive can in fact be seen as responsive to an agent's evaluative conception, and hence as rational action rather than as action brought about by forces external to rational agency. Furthermore, if this account of expressive action underpins an attractive account of emotion, it would also deal with the Redundancy Challenge by showing that rational and emotional explanation of action are not different species: rather the latter would be, in the case of expressive action, a shorthand for the former, indicating that, in acting as they did, the agent was guided by their grasp of the expressive reasons that applied to them in the situation.

My argument begins by looking at an example of action out of emotion that raises the Problem of Expressive Action. I review alternative explanations of the actions in the example and, in seeing why these alternatives fail, I draw up some desiderata for a more adequate explanation. This more adequate explanation will turn out to be the expressive reasons account.

2. Is Action Out of Emotion the Effect of an Affect Program?

Out of grief at the death of Patroclus, Achilles drags Hector's lifeless body repeatedly around the walls of Troy. Patroclus was Achilles' beloved, and it was Hector, prince of Troy, who killed him in battle. Achilles has just managed in turn to kill Hector. Hector's father, King Priam, among others, is watching from the walls above as Achilles, with Hector's corpse, circles below.

The example (from Homer's *Iliad*) is thousands of years old, yet it does not strike us as an unintelligible relic of a defunct culture. Without endorsing Achilles's action, we can find it resonant. We can feel our way into an agential perspective in which that might have seemed the thing to do. In his action we can see wild grief, the dreadful egoism of a man-god, a burning desire for retaliation, and a cold, steely determination that those associated with Hector should understand that this act of revenge will not be Achilles's last.

A temptation, in explaining such a case, might be to assume that emotion has taken Achilles over, that a veil of red mist has fallen over his eyes. However, the *agential insight* that we can have into Achilles's action – the fact that we can understand what it would be like to see that as the thing to do in his situation – suggests that it is unsatisfactory to see Achilles's agency as simply being *overborne* by emotion. To see his agency as being overborne by emotion would be assimilate his action to the class of non-rational behaviour caused by physical or psychological impulses. By contrast, our agential insight into what Achilles does, the resonance of his actions despite their

horror, suggests that we *are* able to appreciate them as performed out of a sense of what counts in their favour.

These considerations count against the explanation of action out of emotion put forward by Kovach and de Lancey (Kovach and de Lancey 2005).⁴ According to these authors, we can explain actions like that of Achilles by appeal to the idea that a class of mental states that they call M-emotions are states with a more or less automatic motivational component that bypasses reasoning and therefore is incompatible with rational explanation. Nevertheless, they claim, explanation of action out of emotion is part of common-sense understanding:

‘Common sense ... tells us that some emotions are states of mind, which are (1) motivational by default, (2) triggered by certain characteristic conditions of elicitation, and (3) cause certain characteristic patterns of response, including full-blown behaviors and actions. What common sense does not tell us is how the emotions possess this type of motivational power.’ (Kovach and de Lancey 2005: p. 108)

If we want a full understanding of why emotions motivate, and why particular emotions motivate as they do, we do best, according to these authors, to look to an account that is based in our best philosophy of biology. The most promising such account, they think, is the ‘affect program’ theory. On this view, the mechanisms that explain action out of emotion are not unique to human beings; rather the explanation proceeds in terms of

⁴ For related views, see (Prinz 2004; Griffiths 1997).

mechanisms that are shared between animals and human beings. Action out of M-emotions is to be explained by reference to an encapsulated 'program' of physiological changes and stereotypical actions that is automatically triggered by certain kinds of stimuli.

It should be noted that affect program approaches don't always take personal-level explanation of actions as falling within their scope. However, Kovach and de Lancey do apply the affect program approach to personal-level explanations. According to Kovach and de Lancey, the affect program view can explain why action out of emotion is not susceptible to 'rationalizing' explanations, where they assume that such explanation proceeds by citing an agent's relevant beliefs and desires. The reason that such explanations give the Wrong Explanation is that emotions often motivate independently of an agent's beliefs and desires. The reason that there is No Further Justification for action out of emotion, and why, as Goldie has it, such actions in the grip of given emotions are primitively intelligible, is that such behaviour is not produced by an agent responding to justifications at all. Rather a program has been triggered that bypasses any such processes of inference.

However, is Kovach and de Lancey's approach a full explanation of actions such as that of Achilles? One problem is that this approach treats Achilles's action as something akin to Anscombe's collector of saucers of mud, or Warren Quinn's turner-on of radios – that is, an action for which there is no normative reason – and implies that we can have genuinely agential insight into it (Anscombe 1957; Quinn 1993).⁵ Yet if there really

⁵ These examples are commonly put forward in defense of the claim that a rational explanation of an action involves citing a feature of the action that the agent reasonably took to be in some way good. By

were No Further Justification of such actions then that would be how we would have to see them. Without relating these actions to the agent's conception of normative reasons, we would not be able to make sense of them in the way that we commonly think we can make sense of a person's actions. Furthermore, emotions do not always – or even often – impel us to act against our better judgement; rather very many cases of action out of emotion are woven unremarkably and apparently seamlessly into our rational life and complement rather than counteract our judgements of what we have most reason to do. We might think that Hume is quite right to say that calm passions accompany much of our experience. They do so without disrupting an intelligible, narratable pattern of activity that takes places in the space of reasons, respects practical norms of inference, and which we can make intersubjectively comprehensible (Helm 2001). Even so-called violent passions can prompt actions that are precisely what we endorse – at the time and on reflection – as that which we have most reason to do in that situation (Srinivasan 2018). Neither can Kovach and de Lancey explain the fact that, in acting out of emotion, we remain *accountable* for what we do: accountable not simply insofar as we are able to impede, temporarily or permanently, the influence of the emotion on our conduct, but for the appropriate and proportionate nature of the emotion and its expression. As R. M. Adams puts it:

‘among states of mind that have intentional objects, the ones for which we are directly responsible are those in which we are responding, consciously or unconsciously, to data that are rich enough to permit a fairly adequate ethical appreciation of the state's intentional object and of the object's place in the fabric

contrast, these examples are said to defy rational explanation because (at least absent some further context) no good seems to lie in these actions.

of personal relationships. Among the states of mind for which we would not be accountable under this criterion are simple feelings of hunger and thirst, insofar as they are primitive responses to physical stimuli, and many states of mind in young children whose experience is not yet rich enough for adequate appreciation of their objects.' (Adams 1985: p. 26)

In Adams's terms, Kovach and de Lancey's view would be an attempt to explain adult action out of emotion as if it were only the effect of simple feelings of hunger or thirst, or of childlike emotions. Presumably these authors appeal to such parallels because they are motivated by an ideal of explanatory parsimony. However, their approach fails to offer a full explanation of adult human action out of emotion. The field of significance to which the emotions of adult human beings respond is far richer, and demands more in the way of developed human capacities, than the stimuli of affect programs.

3. The In Principle Articulatory of Expressive Action

To further substantiate these claims, consider how Kovach and de Lancey's approach neglects the extent to which it is often possible in principle to articulate what the agent saw as good or valuable in action out of emotion. Let us look again at Achilles. We can get a better grip on why Achilles does what he does by seeing that, although his situation is a cause for grief, it is not only grief that is at play. There are further aspects of the situation that are highly salient to Achilles and which help to constitute his perspective on that situation, or his construal of it (Roberts 1988). Achilles is no ordinary person, of course (he is the son of a goddess), and there are few obstacles that he has not been able to beat down by sheer force of will. He is not used to having things going against him, or to meeting unbending resistance. Yet here Patroclus has been

irrevocably taken away from him by Hector. In this situation, his feeling is not simply directed at the loss of Patroclus but also at the fact that the loss of Patroclus represents his (Achilles's) unanswerable defeat by Hector: the loss inflicted can never be undone. Furthermore, so unused is he to defeat that he is simply not prepared to accept it, even in the face of overwhelming evidence. Thus, in response to that defeat he is also defiant: he *will not* be beaten. And so he sets out to inflict a defeat on Achilles that will allow him to emerge in some respect still looking like a victor, that will in part redeem his terrible defeat.⁶ He does that by reducing Hector, his defeater, to less than nothing: reducing him to the status of less-than-human by desecration and by preventing his family from engaging in the necessary rites for their son. Achilles's construal of the situation is something like: how could such a nothing ever be taken to have inflicted a defeat on *me*? His choosing to desecrate Hector's dead body thus shows his grief at the loss of Patroclus, to be sure, but it also shows his rage at having thus been defeated by Hector (that Hector should have been able to injure him, Achilles, in such an inalterable and undeniable way) as well as a certain powerless-yet-defiant attitude that he *will not have it so*.

We can thus get a better grip on why Achilles expresses his grief as he does by bringing into view the highly particular combination of contextual factors to which, as he construes it, he is responding. Even though loss of a beloved is one theme in his situation, there are further salient factors that would call for different kinds of responses – that his situation is experienced as a defeat, say, and that he will not simply

⁶ Cf. the analysis of resentment in (Hampton 1988, p. 57).

rest at being defeated but must come out on top. What Achilles manages to do is to find an action that synthesizes these various factors in a single action.

Achilles's action reflects *his* construal of the situation. A person without Achilles's pig-headed arrogance would not see the situation as he sees it, and would not experience it as calling for the responses that he saw as necessary. We might not expect Patroclus's mother, for instance, to react in the same way to Hector because she would not experience his death as a blow to her pride in the way that Achilles sees it as a blow to his. The way in which Achilles expresses his grief for Patroclus is therefore highly revealing of his evaluative perspective on the situation, precisely because he manages to give such powerful expression to the complex attitudes he has about Patroclus's death. The way he gives expression to his grief is an apt reflection of the complex significance he gives to the death. It is this that allows us to see the deep flaws in Achilles's character embodied in his action, and therefore to see the way in which he doesn't get the situation right at all, or gets some aspects of it wildly out of proportion. As Achilles construes it, the situation becomes all about his own need to demonstrate supremacy over Hector; the fact that Patroclus has lost his life is almost forgotten.

This reading of Achilles's action suggests that Kovach and de Lancey's approach, rather than offering a full explanation of this paradigm case of action out of emotion, is in fact far too simple. Achilles's emotion has a complex intentional content that recognizes various features of his situation as salient; and his action is responsive to that intentional content in such a way that we can read that content from his action by virtue of the fact that he acts rationally on the basis of that content. Thus, we can articulate – and indeed assess – not just the perception of the situation that provides the intentional

content of the emotion, but also the way in which the action is rational on the basis of that perception. This suggests that it would be premature to conclude that Achilles's action is not to be explained as an expression of his rational agency. Rather, the reading we have just given sees Achilles's action as the result of his taking that action to be in key respects good and valuable.

These considerations suggest that rather than there being No Further Justification for such actions – or their simply being primitively intelligible – we can say that an account of actions such as Achilles's should rather respect a condition of *In Principle Articulacy*: that in principle it is possible to articulate both 1) the intentional content of the agent's emotion; and 2) what the agent took to count in favour of the action on the basis of that content. Now whether we *can* indeed articulate the basis for any given action out of emotion is an open question: we need to attempt the articulation and see whether we can get an interpretive story that fits. It is also an open question how far the range of action out of emotion subject to such articulation extends. Perhaps it covers Achilles's case, but what about more simple acts like, say, jumping for joy or ruffling a child's hair out of affection? However, in this section I have established that the articulability of much action out of emotion *is* an open question. Kovach and de Lancey's approach, which treats this question as in principle closed, is not an adequate explanation of these cases; and the No Further Justification objection is misguided.

4. The In Principle Robustness After Reflection of Expressive Action

Nevertheless, it would be misleading to think that we have explained what Achilles does as *instrumentally* rational. Achilles's action cannot really be believed – even by Achilles – *actually* to ameliorate the defeat he has suffered through his loss of Patroclus. If his

action has this aim then it seems to involve a kind of ‘magical thinking:’ Achilles would have to be thought to forget, in the moment of action, that his action cannot possibly bring about the end that counts in favour of doing it.

Indeed, some theorists seem to have thought that this ‘magical’ causality is the source of the difference between expressive and teleological action. For instance, Sartre views emotion as ‘a transformation of the world’, where being in the grip of an emotion allows us to act ‘as if’ something that we know is not possible but nevertheless *want* very much to be possible can be brought about (Sartre 2002: p. 39-40). And while, as we have seen, Peter Goldie takes some emotionally motivated actions to be primitively intelligible, he argues that others can be seen as actions on ‘wishes’, in which the imagination plays a crucial role. On Goldie’s view, expressive action on a wish comes about when the action that is primitively intelligible as an expression of one’s emotion is ruled out by ‘civilizing constraints’ that one accepts. Instead of performing that forbidden action, one rather performs an action that is ‘symbolically related’ to the primitively intelligible action that one desires to do. Goldie takes this as a potentially good explanation of emotional actions such as gouging out the eyes of a picture of one’s rival, or smashing one’s spouse’s favourite vase in the middle of a blazing row (Goldie 2000).⁷

⁷ Kovach and De Lancey, by contrast, offer to explain such cases, not in terms of the imagined or symbolic fulfilment of purposes, but rather in terms of the phenomenon of ‘redirection’ of emotion, which they take to be exhibited in animals as well as humans, and thus not to require the powers of symbolic imagination appealed to in Goldie’s account. For instance, when a lioness whose tail is bitten by a cub takes it out on a nearby tree trunk rather than the cub itself, she is not doing it because the tree trunk is a symbolic substitute for the cub, but rather (presumably?) because she has a strong urge that she needs to release, and must do so without damaging to her cubs. However, as noted above, this apparently more parsimonious explanation in terms of responses humans share with non-human animals is not necessarily a better one if it fails to capture the richness of the structure of adult human emotions.

The problem with these explanations, however, is that the examples of expressions of emotion from which they generalize are *inappropriate*, in a sense I will now explain; and that the analysis of expressive action that they offer cannot easily be transferred to those cases of *appropriate* expression of emotion, which are significantly more numerous. In considering Achilles's action, for instance, we need to bear in mind that it is an *inappropriate* expression of his grief. If there is something that Achilles takes to count in favour of acting in this way given what has happened to him (and to Patroclus), it does not *actually* count in its favour. This is not simply to say that his action is morally inappropriate. Rather it is the kind of action of which we might say, 'I know you are grieving, but how could you do *that*?' However, there presumably are *appropriate* expressions of grief. Indeed, it seems an engrained part of the common-sense of action-explanation that there are appropriate ways for emotion to influence conduct, and that relevant examples, such as admirable expressions of sympathy, of concern, of grief, of anger, of resentment, of remorse, of shame, are ubiquitous. Appropriate expressions of emotion are also central to what feminists have called *emotional care work*: those interactions in which the support that one offers to another comes about in part through the emotions that, in response to their situation, one experiences and displays. (Ruffling a child's hair out of affection might be part of such care work.) An account of action out of emotion should be able to explain cases such as Achilles, but it should also be able to explain cases of appropriate action out of emotion.

The 'magical thinking' account, however, does not give a convincing explanation of such mundane but important forms of action out of emotion as embracing someone out of sympathy, or because one is glad to see them, or sorry to see them go. Such actions do not seem to be the wished-for or symbolic fulfilment of a primitively intelligible

purpose that civilization denies us. Sartre might say that our embrace of sympathy is based on magically thinking that one can look after the person and solve all of their problems. Or that the embrace of leave-taking is based on magically denying that one does in the end have to let the person go. However, the problem with 'magical thinking' explanations is that if they were true, they would be undermined by reflection on their true nature. Once the agent became aware that their emotional action was founded on a delusion, it would put them at odds with their own action: their action would, on reflection, appear to them, not merely the arational irruption of a force external to their agency, but as a form of irrationality taken up and pursued by their agency. However, this would be a highly counter-intuitive way of seeing, e.g., emotional care work. On the face of it, there are many expressive actions that do not strike us as being irrational on reflection. Thus while the 'readings' of various emotions that Sartre and Goldie offer can be illuminating and insightful, they cannot be the whole story about the intelligibility of action out of emotion. The 'magical thinking' account is too sweeping: it explains all action out of emotion on the basis of some eye-catching but inappropriate examples.

We can therefore give a further desideratum for a more adequate explanation of action out of emotion: that it allows at least some, and perhaps many, such actions to be *In Principle Robust After Reflection*. The In Principle Robustness After Reflection condition adds the further desideratum that an adequate understanding of cases such as Achilles's should not rule out in principle that our sense of expressive action being the thing to do when in the grip of an emotion might survive reflective awareness of the nature of that action-tendency. In other words, we want it to be possible that in at least some – and perhaps many – cases of emotional action we can know exactly what we are doing but still think it the thing to do.

5. The Expressive Reasons Account

Rather than ascribing a magical belief to agents acting out of emotion, I suggest that we rather see them as, at least potentially, acting in ways that they take themselves to have good normative reason to act. As I have noted, many actions out of emotion appear to cohere with actions motivated by our evaluative conception. However, the Problem of Expressive Action that I noted at the start points out that many such actions cannot be seen as guided by those normative reasons recognized by standard views of practical rationality. If, as I have suggested, it would be unacceptable not to recognize the rationality of such action out of emotion, we rather need to consider expanding the range of normative practical reasons that, as theorists, we recognize.

The explanation I will put forward focuses on a category of *expressive actions*; and action on *expressive normative reasons*. Goldie is on the right lines, I think, in appealing to the symbolism of expressive actions. However, he goes wrong in assuming that the only way that action out of emotion could be symbolic would be as the wishful diversion of a primitively intelligible but impermissible action. As we have seen, this proposal fails the test of In Principle Robustness After Reflection. However, perhaps there is a more positive role for symbolism, on which it is possible that symbolic actions do pass the Robustness After Reflection test.

Sometimes what is important to people is to find a response to a situation that does not attempt to change it but rather that acknowledges or does justice to its *gravity*, or more generally, what is salient in it (Bennett 2016). Take, for instance, the case of a religious believer, overcome with emotion in the presence of their god, falling to their knees in

church. Following the symbolic action approach, we could say that this action symbolizes the attitude of reverence and awe that believers feel in the presence of a being whose worth is incomparably higher than their own. It is an appropriate symbol because lowering oneself reflects the highness of one's god. However, assuming for a moment that this is plausible, why do they seek thus to symbolize their attitudes? My suggestion is that in doing so they are seeking to *do justice* to the situation as it strikes them. Their action attempts in some adequate way to mark it, or acknowledge its importance.

The thought here is that there are simply certain moments in our lives that we need to recognize as special, to mark in some adequate way, either to celebrate or deprecate. In attempting to do justice to such situations, we thus dwell on events and we lift them out of the interminable onrush of indiscriminate events ('one damn thing after another') and give them a special status in our lives. Dwelling on them is part of taking them seriously. These moments may be moments of significant achievement; or of wrongdoing or evil; moments of rupture or repair; moments of significant social or personal change. Or they may simply be moments in which some important aspect of reality is disclosed to one's view, as with the way in which one's child suddenly appears to one as particularly charming, or cute, or affection-worthy – or whatever is captured by the action of ruffling her hair (and there are many different ways that this action can be performed, corresponding to different aspects of the child being perceived as salient). In lifting these moments out of the mundane, we also create memories of them that will last – and this has value both for what we can learn from them, but also because some things demand not to be forgotten in the way most of the moments of our lives are quickly forgotten. This helps to ensure that these moments will be given the

special treatment they deserve, but also make our own lives more meaningful, since the tenor of our lives is not merely instantaneous but has peaks and troughs and a thread of connections to the past.

If, as I argue here, we do sometimes seek to do justice to significant situations in which we find ourselves, we need some basis on which to do so. Our starting point – it is hard to see where else we might begin – is a vocabulary of *expressively powerful* actions, actions that, by virtue of symbolic connections with the nature of the situation to which they refer, can be understood as a satisfying way to capture or distil the significance of the situation. This is how we can understand the kneeling of the believer; or the self-covering of the person gripped by shame; or the distancing from the wrongdoer enacted by the person righteously condemning; or the embrace of the person filled with sympathy for a suffering other. Otherwise put, action out of emotion expresses an emotion by virtue of possessing expressive properties that relate appropriately to the situation to which the emotion is directed. If these expressive properties count in favour of an action as a way of doing justice to the situation then we can say that we have expressive reason to act in that way. If the best explanation of why we act in that way is that we take ourselves to have reason so to do justice to the situation in virtue of the expressive properties of the action then we can say that we act on expressive reasons.

We can therefore distinguish between acts that are ‘expressive of’ some attitude or emotion, in the sense that I have in mind, and acts that are caused by some attitude or emotion without being ‘expressive of’ it. Take being caused to run away by the sight of a human shape unexpectedly visible in the corner of your room as you return to your dark house one night. Your act of running away may be caused by your fear. But it is not

expressive of it in the sense that I am interested in. We can see actions that are 'expressive of' an emotion as being taken by their agents to be appropriate by virtue of being *expressively powerful* in relation to the situation as the agent experiences it. We perform such actions guided by a sense, in other words, of the power of these actions as a response to what is salient in the situation as we construe it.

Expressive reasons are a *sui generis* category of normative reason, different from instrumental reasons, prudential reasons, reasons of obligation or consequentialist reasons. A full specification of an expressive reason for any particular action would be as follows:

- a) In at least some situations, there are reasons, not just to promote a further end, but rather to do justice to the import of the situation.
- b) In order to do justice to the situation one has reason to do something that is symbolically related to the situation.
- c) There are reasonably determinate criteria of symbolic adequacy, so one has reason to do an action of a particular sort when one has the aim of doing justice to one's situation.

If this analysis is plausible then the Problem of Expressive Action would be solved. The solution to that problem would also show that the Wrong Explanation challenge fails, as does the No Further Justification challenge. Against the No Further Justification challenge, we have seen that desiderata on an explanation of expressive action are its In Principle Articulacy, and its In Principle Robustness After Reflection. Assuming that expressive action is action out of emotion then the Redundancy challenge also fails.

It might be objected, however, that my response to the problem of the rationality of action out of emotion is in terms of expressive action, whereas it is possible to engage in expressive action without emotion. An action is expressive, I have claimed, by virtue its expressive properties. But what makes these properties expressive is their relation to features of the situation. In other words, an agent who, say, kneels in church may be doing something that has powerful expressive properties, even if the agent themselves are not themselves particularly moved by the situation. If this is correct then the agent's action might have the relevant expressive properties, and thus be expressive of awe, even though it does not express *their* awe. It might thus be argued that the expressive reasons account has not explained the target phenomenon: that is, those actions out of emotion that raise the Problem of Expressive Action. Expressive action (i.e. action with expressive properties) is not the same as action out of emotion.

However, this conclusion would be unwarranted. It is true that it is possible to engage in expressive action without emotion. But it is possible to act in all sorts of ways characteristic of emotion without experiencing that emotion – and without the action being done from the emotion. Nevertheless, this objection raises some pertinent questions. If expressive action involves acting on expressive reasons, and much action out of emotion is expressive action, how is acting on expressive reasons out of emotion different from acting on expressive reasons without emotion? How does an awareness of expressive reasons contribute to the development and individuation of emotions?

Now it is true that I have not given a full theory of the emotions – not even of those emotions for which the characteristic actions can be explained as expressive actions.

And it might seem strange to have a theory of action out of emotion without having a theory of emotion. However, it is likely that a number of theories of emotion might be compatible with what I have to say about expressive action. Nevertheless, we can ask what would have to be true of the emotions for the expressive reasons account to be a good explanation of (some) action out of emotion can be compatible with canons of rational explanation. In answering this question, we should bear in mind that we are looking to capture the conception of emotion to which we appeal when, in explaining actions, we adopt the perspective of what P. F. Strawson calls 'ordinary, adult interpersonal relationships' (Strawson 1962). According to the Strawsonian approach, the 'interpersonal' understanding we are talking about is of a fundamentally different order to that we achieve when we explain a part of nature as governed by causal laws. For instance, were it possible to understand a highly complex phenomenon non-rationally, such as the weather system in a given locality, this would involve understanding causal mechanisms that work in a law-like way to produce certain kinds of weather given certain background conditions. By contrast, the understanding we seek in the context of 'ordinary, adult interpersonal relationships' explains what happens largely as the outcome of *agents responding to what they rationally take to be normative reasons*. Thus in saying what an emotion is within the interpersonal perspective on human behaviour, we should not take for granted accounts of the emotions, grounded in the natural sciences, that leave rational agency out of the picture. Indeed, I have tried to show some of what would be lost if we did not see action out of emotion as action governed by the agent's conception of normative reasons.

Beyond that negative point, what the theory of expressive action helps to explain is the nature of the connection between (i) the *object* of the emotion – the nature of the

situation to which the emotion responds – and (ii) the *motivational tendency* of the emotion – that is, the pattern of action to which the emotion characteristically prompts. An emotion is in part a regular pattern of action-readiness and execution, accompanied by feelings associated with arousal to act, and patterns of attention that tend to focus on the aspects of the situation to which those actions are geared to respond. One crucial question about emotion concerns the nature of this object-action link, and hence in virtue of what the various elements of an emotional response hang together. The expressive action theory says that the object-action link of the relevant kind of emotion makes a claim within the space of reasons: the object-action link effectively says that a given situation contains something of value or disvalue that needs to be lifted out of the ordinary run of events and marked as significant; and that *this* action is the appropriate way to respond, or to do justice to, that particular value or disvalue.

Now this line of thought might suggest a particular story about the emotions. For instance, we might speculate that expressive actions arise at some point in human cultural development. The explanation for this might be the nature of a particular culture, or our inherited emotional psychology, or our engagement with an objective realm of value, or some hybrid of all three, depending on which historical and meta-ethical stories are the right ones. Further, imagine that, through patterns of socialization, two things happen. Firstly, perhaps through their adoption in shared forms of ritual, particular expressive action types become socially entrenched, thereby cementing certain object-action links in the collective consciousness. And secondly (as a result of socialization, or our inherited psychology, or our engagement with the realm of objective value), people yet come to find such acts satisfying vehicles through which to express their own emotions. In that case we would have an explanation of how certain

emotions (emotions that are socially constructed, in at least some sense of that term) would involve object-response pairings that are widely shared and found expressively powerful.

This story is speculative, but if it is along the right lines, we would have an explanation of how expressive action explains (some) action out of emotion even if expressive action does not require emotion. It would explain such action out of emotion because expressive action figures in the full explanation of the development and individuation of the emotions. Our story about expressive action would explain how the range of ways in which action out of emotion can be guided by normative reasons expands beyond prudential, instrumental, deontic or consequentialist reasons. Furthermore, there would then be a ready explanation of why much expressive action is non-accidentally done out of emotion: for instance, when it involves a heightened, engaged awareness of the force of those reasons.⁸ The expressive action account would also explain how the various components of the relevant emotions hang together as a coherent whole. The intentional object of one of the relevant emotions, for instance, would be that aspect of the situation that is valuable in precisely such a way as to call for the prompted action. Emotion phenomenology would be explained through attentional focus on that particular value or disvalue while the process of arousal towards the associated action is going on.

If some such story is true, it would also help to address the Redundancy objection by showing how explanations of action in terms of following expressive reasons are

⁸ This is an important point of overlap between the approach of the present paper and that of (Carman 2018a).

entangled with explanation in terms of emotion. Expressive reasons would not only be reasons for action but would also be reasons for emotion. Furthermore, it would explain how reasons for action and reasons for emotion are related to one another: if one has a reason to experience an emotion one would, if the appropriateness of the emotion is based on the appropriateness of the expressive action, thereby have at least a *pro tanto* reason for action, at least in the case of those emotions with a significant motivational component.

Nevertheless, it might seem implausible that in acting spontaneously out of emotion we could really be thought of as doing something as elaborate as symbolizing the intentional content of our emotions, and thereby following expressive reasons. Have I been too quick with the idea that action out of emotion is rooted in the non-rational? Now of course all of our rational capacities are to some degree rooted in the non-rational: only because there is much of us about us that is non-agential can we be agents. We do not need to deny that action out of emotion is rooted in what is not rational, even if we insist that it can be guided by normative reasons. The point for which I have been arguing here is that, with respect to a significant class of action out of emotion, we will not have a full explanation of the action by pointing *only* to non-rational processes and mechanisms. We will not have a full explanation until we understand the reasons on which such actions are done.

Furthermore, the standpoint of rational explanation does not rule out spontaneous action. Its reference to the agent being guided by a conception of normative reasons does not rule out e.g., a highly skilled football player's manoeuvres in the fast rough-and-tumble of an unfolding game. Rational explanation can also accommodate habitual

actions such as successfully following the route to work while having a complex conversation that takes all of one's attention. In these cases, the actions can still be seen as rationally guided to the extent that they are controlled by – and would not have been performed without the existence of – an overarching grasp of what one has reason to do.⁹ If expressive action along the lines sketched here were such that we learned from an early age what kinds of acts were expressively appropriate to what broad range of situations, then there seems to be no reason to think our vocabulary of such acts could not become as spontaneous as our use of language, or of other skilled and habitual practices into which we are inculcated from birth. In Achilles's situation the symbolism of annihilating Hector comes naturally to him, given his inheritance of a certain vocabulary of expressive action and his application of it to his situation. This is a matter of second rather than first nature.¹⁰

Nevertheless, there might be a worry that, if my story were correct, it would have to appear more obvious to us than it does that when we act out of emotion, we are in fact seeking to do justice to our situation. However, the symbolic action approach doesn't have to insist that e.g. in ruffling my child's hair I act, at that moment, under the description, 'doing justice to my situation,' or 'symbolizing my attitudes'. Sometimes we can have an adequate grasp of our reasons without having a complete grasp.

Furthermore, another possibility might be that in performing the actions that I have called expressive, we do not always, or even often, follow expressive reasons. We might simply perform these actions because we have learned that this is the socially accepted

⁹ Cf. the discussion of 'intentions-in-action' in (Searle 1983: pp. 84-5). For the notion of action being controlled by a purpose, see e.g. (Pettit 2018). In respect to action out of emotion, see for instance the discussion of executive intentions in (Pacherie 2002).

¹⁰ To use John McDowell's distinction. See (McDowell 1994).

'script' for these kinds of situations. However, in discovering the reason for the action – as supplied, let us assume, by the expressive reason account – one would gain a greater insight into one's own action, and would be enabled to act rationally out of emotion in a way one had not previously been able to. The expressive reason account would thus have the effect of allowing us to achieve greater transparency over the rationality immanent in the social forms that, up to now perhaps, we may simply have blindly adopted. Increasing transparency in regard to our reasons is an important ideal: it puts us in a position of greater rational control over our action, and thus allows our actions more perfectly to approximate the ideal of rational thought translated into action. This is a kind of Kantian or Hegelian line of thought. Greater rational control means greater self-determination, and a gradual overcoming of the heteronomy of being the unreflective agents of nature or society. Alternatively, of course, we might come to realize that these emotions are based on object-action links that became ingrained in our psychology as a result of rituals from a now-distant culture, and that once articulated, which we can no longer find resonant. Either way, articulating the expressive reasons that guide much of our action out of emotion promises greater self-understanding and greater rational control.

6. Conclusion

At the outset of the paper, we saw that rational explanation of action out of emotion faces a number of challenges. The Wrong Explanation Challenge says that explaining action out of emotion by reference to a purpose rather than an emotion gets it wrong. The Redundancy Challenge says that if explanation of an action by reference to emotion is sufficient then rational explanation is redundant. And the No Further Justification Challenge says that there is no more to say, at the level of rational explanation, about

why people act as they do out of a particular emotion. Furthermore, even if these challenges can be addressed, there is a Problem of Expressive Action, since many actions out of emotion seem unpromising candidates for being guided by normative practical reasons of the prudential, instrumental, deontic or consequentialist sort. However, in thinking about how to explain the expression of Achilles's grief and anger, I have argued that we have to see his action, and that of appropriate expressions of grief and anger, as expressive actions guided by a conception of normative practical reasons rather than as either the effect of an affect program or the symbolic fulfilment of a frustrated wish. Expressive action is guided by the agent's conception of expressive reasons, insofar as the agent seeks to *do justice* to the situation by virtue of the action's *expressive power*. Action out of emotion can be seen as responsive to an agent's evaluative conception, and hence as rational action rather than the product of forces external to rational agency.

REFERENCES

R. M. Adams, 'Involuntary Sins,' *Philosophical Review* 94 (1985)

Maria Alvarez, *Kinds of Reasons: An Essay in the Philosophy of Action* (Oxford: Oxford University Press, 2010)

Elizabeth Anderson, *Value in Ethics and Economics* (Cambridge MA: Harvard University Press, 1995)

Elizabeth Anderson and Richard Pildes, 'Expressive Theories of Law: A General Restatement,' *University of Pennsylvania Law Review* 148 (2000), 1503-

G. E. M. Anscombe, *Intention* (Oxford: Blackwell, 1957)

Christopher Bennett, 'Excuses, Justifications and the Normativity of Expressive Behaviour,' *Oxford Journal of Legal Studies* 32 (2012), 563-581

Christopher Bennett, 'Expressive Actions' in Catherine Abell and Joel Smith (eds), *The Expression of Emotion* (Cambridge: Cambridge University Press, 2016), 73-94

Jason Brennan and Peter Jaworski, 'Markets Without Symbolic Limits,' *Ethics* 125 (2015), 1053-1077

Mary Carman, 'Emotionally Guiding Our Actions,' *Canadian Journal of Philosophy* 48 (2018a), 43-64

Mary Carman, 'How Emotions Do Not Provide Reasons to Act,' *Philosophia* 46 (2018b), pp. 555-574

Justin D'Arms and Dan Jacobson, 'The Moralistic Fallacy: On the "Appropriateness" of Emotions,' *Philosophy and Phenomenological Research* 61 (2000), 65-90

Peter Goldie, 'Explaining Expressions of Emotion' *Mind* 109 (2000), 25-38

Patricia Greenspan, *Emotions and Reasons: An Enquiry Into Emotional Justification* (New York: Routledge, 1988)

Patricia Greenspan, 'Emotional Strategies and Rationality,' *Ethics* 110 (2000), 469-487

Bennett Helm, *Emotional Reason: Deliberation, Motivation and the Nature of Value*
(Cambridge: Cambridge University Press, 2001)

Paul Griffiths, *What Emotions Really Are* (London: University of Chicago Press, 1997)

J. Hampton, 'Forgiveness, Resentment and Hatred' in J. G. Murphy and J. Hampton,
Forgiveness and Mercy (Cambridge: Cambridge University Press, 1988), 35-87

R. Hursthouse, 'Arational Actions,' *Journal of Philosophy* 88 (1991),

Adam Kovach and Craig De Lancey, 'On Emotions and the Explanation of Behavior,' *Noûs*
39 (2005), 106-122

John McDowell, *Mind and World* (Cambridge MA: Harvard University Press, 1994)

Elisabeth Pacherie, 'The Role of Emotions in the Explanation of Action,' *European
Review of Philosophy* 5 (2002), pp. 53-91

Philip Pettit, 'Three Mistakes About Doing Good (and Bad),' *Journal of Applied
Philosophy* 35 (2018), pp. 1-25

J. Prinz, *Gut Reactions: A Perceptual Theory of Emotion* (Oxford: Oxford University Press, 2004)

Warren Quinn, 'Putting Rationality in Its Place' in Quinn, *Morality and Action* (Cambridge: Cambridge University Press, 1993)

Robert C. Roberts, 'What An Emotion Is: A Sketch,' *Philosophical Review* 97 (1988), 183-209

Jean-Paul Sartre, *Sketch for a Theory of the Emotions*, trans. P. Mairet (New York: Routledge, 2002)

T. M. Scanlon, *What We Owe To Each Other* (Cambridge MA: Belknap Press, 1999)

John Searle, *Intentionality* (Cambridge: Cambridge University Press, 1983)

John Skorupski, *The Domain of Reasons* (Oxford: Oxford University Press, 2010)

Amia Srinivasan, 'The Aptness of Anger,' *Journal of Political Philosophy* 26 (2018), 123-144

Angela Smith, 'Responsibility for Attitudes: Activity and Passivity in Mental Life,' *Ethics* 115 (2005), 236-71

P. F. Strawson, 'Freedom and Resentment,' *Proceedings of the British Academy* 48 (1962), pp. 1-25

CHRISTOPHER BENNETT is Professor of Philosophy at the University of Sheffield. His recent publications include 'The Alteration Thesis: Forgiveness as a Normative Power,' *Philosophy and Public Affairs* 46 (2018); and 'The Authority of Moral Oversight: On the Legitimacy of Criminal Law,' *Legal Theory* 25 (2019).