

WHAT GOES ON WHEN WE APOLOGIZE?

Christopher Bennett

APOLOGY is often said to play an important role in reconciliation. On a plausible interpretation of that claim, apology has this important role because the performance of an apology provides us with new practical reasons, reasons to change the way we relate to the wrongdoer. But if this is right, what kinds of reasons are they, and why is an apology necessary (and sufficient) to provide us with such reasons?

In this paper, I argue that our practice of giving and demanding apologies is underpinned by a belief that apologies make a difference to our normative situation: that once an apology has been given, the rights and responsibilities of the apologizer and others have been altered. However, if we ask what rationalizes that belief, two influential views in the literature on apology—which I call the *reassurance view* and the *performing deference view*—prove to be inadequate. One thing a theory of apology needs to explain is that the distinctive work of reconciliation carried out by apology involves a set of canonical actions through which one can change one's normative situation in characteristic ways. However, it is also a characteristic feature of apology that (at least in cases of serious wrongdoing) it effects this reconciliation only when sincere—that is, when it is an expression of the wrongdoer's remorseful recognition of the wrongness of what was done.¹ The reassurance view and the performing deference view fail to offer an explanation that is adequate to both of these features. In order to explain these features of apology, I suggest that we see apology as a power to change one's normative situation through the performance of canonical actions, but a power that is exercised expressively, or by an *expressive action*.²

For an apology to do its distinctive work, I will argue, it needs to be an expression of emotion that is *appropriate*. This means not only that the *emotion* should be appropriate but also that the emotion should find an appropriate

1 Govier and Verwoerd, "The Promise and Pitfalls of Apology."

2 In offering an account of apology as an expressive action, I will be drawing on an account of expressive action I have developed in a number of other recent papers. See Bennett, "Expressive Actions," "The Problem of Expressive Action," and "How and Why to Express the Emotions."

expressive *vehicle*. This, it will turn out, means that the expression of the relevant emotion has to conform to certain canonical features. Once we understand apology as an expressive action, we can see how, in favorable conditions at least, the canonical actions through which the apologizer's normative situation can be altered could also be those through which individuals can give authentic expression to their emotion. In this way, we can explain how it might make sense to see apologies as providing new practical reasons while at the same time expressing the wrongdoer's authentic remorse.

I

There are many things it is appropriate for a wrongdoer to do after wrongdoing. What I am interested in is understanding the normative work done specifically by apology. We can focus on this question by asking what is missing when an apology fails to be given. Imagine someone saying the following:

Missing Apology: I know that he's sorry for what he has done. And I am sure that he wouldn't do it again. I trust him and don't want to lose him as a friend. In fact, I think that despite everything that happened we *are* still good friends. I know he has changed since then, and he knows that I know it too. It just feels a bit strange that he has never apologized. It has just left everything a bit unresolved. Like there is unfinished business. And somehow I just can't feel entirely good about my relationship with him anymore because of it. I am sure he really does feel bad about what he did to me. But why won't he just come out with it?

The speaker in this example is insisting that *an apology* is what they are owed and that until an apology is forthcoming, matters are not settled. This is clearly not to say that only apology matters. The speaker is grateful that their friendship with the wrongdoer has survived. And they have an appreciative attitude to whatever the wrongdoer has done to rebuild trust between them. However, an apology has not been given, and this has left our speaker with the perhaps hard-to-pin-down feeling that something remains unfinished. In this section, I attempt to articulate what lies behind the speaker's feeling that there is unfinished business where an apology is not forthcoming.

To start with, let us get clear on what apologizing involves. I am interested in what we might think of as an unreserved apology.³ Looking at what is involved in apologizing unreservedly—where the apology is sincerely given and

3 See Davis, "On Apologies," on the "consummate apology." See also Smith, "The Categorical Apology."

unconditional—is relevant here because it can help us to see what elements are required for an apology to leave no business unfinished. In the following paragraph, I aim simply to summarize the features standardly given in accounts of such full, unreserved apologies.⁴

First of all, such apologizing normally involves speech addressed to the wronged party—an “I’m sorry.” However, as is often said, sorry is not good enough; so, second, this speech should involve an acknowledgment of responsible wrongdoing (“It was my fault”) and, third, a credible commitment to refrain from such acts in the future (“I can see that it was wrong, and I won’t do it again”). Fourth, for cases of nontrivial wrongdoing, apologizing unreservedly involves the wrongdoer showing that they are troubled by the wrong they have done to their victim; the apology is thus an expression of remorse. Apologies for nontrivial matters are undermined by the appearance that the apologizer does not feel remorseful for what they have done.⁵ Fifth, apology will also usually involve an offer of restitution and a commitment to make amends. Sixth, a person who is apologizing unreservedly does not do so stridently, confidently, unabashedly; rather their demeanor, posture, and gestures exhibit deference and humility. Seventh, a successful apology involves some credible attempt to make the extent of the remorse—the amends offered and the degree of humility—proportional to the perceived seriousness of the wrong. The wrongdoer thinking that an admission of responsibility and a commitment to refrain in the future are sufficient, or offering to make amends but only to a negligible degree, can reveal that they underestimate the seriousness of the wrong. An apology can thus misfire if it fails to include one or more of these seven elements (at least when it could reasonably have included them) or if what is offered fails to be proportionate to the seriousness of the wrong.

So, how can we start to articulate what lies behind the view expressed in Missing Apology? First of all, Missing Apology seems to imply that among appropriate responses to wrongdoing, it is reasonable to give and ask for apologies in particular. As we have noted, this is not to say that apology is the only appropriate response to wrongdoing. It is not even to say that it is the most

4 In addition to the papers already cited, see, e.g., Tavuchis, *Mea Culpa*; Joyce, “Apologizing”; Gill, “The Moral Functions of an Apology”; Lazare, *On Apology*; Smith, *I Was Wrong*; Bovens, “Apologies”; Bennett, *The Apology Ritual*; Martin, “Owning Up and Lowering Down”; Pettigrove and Collins, “Apologizing for Who I Am”; and Helmreich, “The Apologetic Stance.” In defending the possibility of vicarious and collective apologies, Andrew Cohen argues that apologies should be thought of in terms of characteristic functions rather than required features (see “Vicarious Apologies as Moral Repair”).

5 For some discussion of counterexamples that involve actions for which one says sorry but that were ultimately beneficial and thus not, it is claimed, cause for regret, see Barnum-Roberts, “Apologizing without Regret.”

important one. But it is to say that a wrongdoer's response to wrongdoing is incomplete if an apology is missing. It seems to be a corollary of this that the giving of an apology makes a moral difference that could not have been brought about by other means. However, if we want to say precisely what this moral difference is, we should note that Missing Apology also implies that while apology has a role in bringing about reconciliation after wrongdoing, it is neither necessary nor sufficient for actual resumption of good relations. Apology is not *necessary* for good relations since it is realistic to think that although the speaker is aware that something is missing, the pair in Missing Apology are getting along *okay*, with trust and goodwill and friendship. And apology is not *sufficient* for good relations because apology cannot bring about good relations in the absence of whatever additional trust-, friendship-, and goodwill-(re)building measures the wrongdoer in this case has undertaken. Apology cannot bring trust about magically, just by the uttering of "I am sorry; it's my fault; I won't do it again." Despite this, the speaker implies that apology is necessary and sufficient for *some* kind of closure after wrongdoing, since they think that some sort of closure is lacking because an apology has not been given.

I suggest that we look for what is distinctive in apology, not in its role in bringing about actual good relations, but rather in the way that it alters the normative situation that arises from wrongdoing. We can interpret the dissatisfaction expressed in Missing Apology as evincing an awareness that even though he may have done many other things that are appropriate, the wrongdoer has not brought about some alteration of the normative situation that apology can (and should) bring about. The speaker is in two minds because they recognize the good in the wrongdoer, and the trust that has been rebuilt, but nevertheless feel dissatisfied because there is an element of moral compromise in resuming relations in the absence of an apology, going forward as if everything were normal. Even though he has rebuilt confidence, the wrongdoer has not done that specific thing that would allow our speaker, in clear conscience and while doing full justice to the significance of the wrong, to resume normal relations with him. The speaker feels themselves to be in a situation of moral compromise because of the lack of an apology.

Now, in trying to make sense of Missing Apology, it is crucial to note that it is possible, as I have done, to specify *in advance* of any particular apology the elements that such an apology will need to involve in order to bring about the relevant kind of closure. As moral agents familiar with the normative expectations involved in giving and demanding apologies, we know in advance that apologies that are not addressed to the victim (when they easily could be), that do not involve a commitment to refrain in the future, that are not appropriately remorseful and deferential and proportionate, etc. will not cut the mustard.

This suggests an important explanandum for a theory of apology: that, whatever moral work is done by an apology, the elements required to make it normatively effective are relatively unchanging across contexts. Whereas rebuilding trust, or salvaging a friendship, will require a lot of contextual information about the person involved, the nature of the relationship, and the extent of the hurt feelings caused by the wrong, apologies do not work like that. While of course there is some room for subjective variation, it is possible to say in advance that an apology that does not contain the relevant elements (when it reasonably could have) will misfire, failing to bring about its characteristic normative effects. The appropriateness of the elements of apology to cases of wrongdoing is thus to an important degree independent of context. I will refer to this feature by saying that an unreserved apology is made up of a set of elements that are *canonical*.

These considerations count against a theory of apology that I will call the *reassurance view*. According to the reassurance view, apology gives us new practical reasons by providing evidence of psychological change in the wrongdoer. If it is often said that wrongdoing ruptures relationships, one version of the reassurance view sees this rupture as damaging the victim's confidence in their *own* moral standing. Another version, not necessarily incompatible with the first, sees it as damaging confidence or trust in the *wrongdoer*. The central feature of the reassurance view is thus that it sees wrongdoing as causing some harm, and apology as repairing that harm. Apology can repair the harm caused by wrongdoing, on the reassurance view, by the wrongdoer demonstrating a renewed commitment to moral standards. The distinctive moral effect of apology therefore consists in giving credible and practically relevant *evidence* of psychological change in the wrongdoer.⁶

However, as we can see from the preceding discussion, we can readily imagine a situation, like Missing Apology, in which the wrongdoer's actions have been such as to allow the victim to regain their confidence in their own moral standing and to allow the victim and others to place their trust in the wrongdoer again, but where no apology has been forthcoming. In such a situation, I

6 For an influential source of the reassurance view, see Murphy and Hampton, *Forgiveness and Mercy*. See also Gill, "The Moral Functions of an Apology." The reassurance view of apology might also fit well with Margaret Urban Walker's account of "moral repair" as "restoring or creating trust and hope in a shared sense of value and responsibility" (*Moral Repair*, 28). Similarly, Adrienne Martin puts it thus: "An apology typically includes saying one knows one flouted a legitimate norm and regrets it; the recipient needs to know the wrongdoer understands that he acted from an inadequate interpersonal commitment, if the recipient is to have reason to cease resenting that inadequate commitment" ("Owning Up and Lowering Down," 28). Note, however, that on Martin's account apology has performative as well as reassurance elements.

have suggested, it makes sense to think that there is normative work for apology to do that has not yet been done. However, the reassurance view does not have the resources to account for this. The job that, according to the reassurance view, it is the place of apology to carry out has in *Missing Apology* already been done.

Furthermore, I have argued that the elements required for a satisfying apology are canonical and largely determined independently of context, whereas if the reassurance view were correct, the acts that will provide the kind of evidence of psychological change a particular audience will need in order to be appropriately reassured would depend very much on the particular wrongdoer, the nature of the relationship, the particular victim, the expectations and biases of the audience, etc. Indeed, since evidence can be better or worse, the reassurance view gives us an account of the normative role of apology that is context dependent and scalar, thus failing to capture what is distinctive about the all-at-once nature of the closure hoped for in *Missing Apology*. A proponent of the reassurance view might argue, in rule-consequentialist fashion, that we have settled on a canonical set of elements as a kind of shorthand for signaling genuine remorse. But then we might query why performance of the shorthand would serve as good evidence of *genuine* remorse, rather than just a willingness to signal such remorse. At the very least, we would need a story about how the canonical elements could come to be the vehicle for genuine expressions of emotion.⁷

In trying to show that apology has an important normative function, the reassurance view takes it that this function is to repair the *harm* done by wrongdoing. More plausible, however, if we want to explain the distinctive role of apology in moral repair, is that the moral function of apology lies in addressing the fact that the victim has been treated *wrongfully*.⁸ The attractions of the reassurance view are clear, since repairing the harms of wrongdoing is indeed an important job and we can see in clear, nonmetaphorical terms how trust and confidence could be rebuilt after wrongdoing (albeit that it might be hard to do). By contrast, the idea of addressing past wrongs might seem less immediately urgent a task and perhaps even hopelessly metaphorical. However, if we want to explain why apology has a distinctive role among responses to wrongdoing, I will argue that we need to explore the idea that its function is to act on the normative situation *directly* and not simply as a source of evidence of the wrongdoer's state of mind.

7 For one version of such a story, see Pettigrove and Collins, "Apologizing for Who I Am," 144–48.

8 Cf. Hampton, "Correcting Harms versus Righting Wrongs."

II

In search of such an alternative, then, we might turn to the *performing deference view* of apology. The performing deference view understands the situation of wrongdoing as one in which the wrongdoer has subjugated the victim, degrading or demeaning them by treating them as lacking in moral status and thus as one whom it was permissible to treat in that way. This act creates harmful psychological effects, and, more saliently for the purposes of apology, it is a wrong done by the perpetrator against the victim. More metaphorically, perhaps, it is a disturbance in the normative order: an act that contravenes basic requirements of respect and consideration. On the performing deference view, the way to address the wrong and undo that normative disturbance is for the wrongdoer effectively to reverse that situation and to act out their subordination to the victim by means of apology. As Jeffrie Murphy puts it:

Wrongdoers attempt (sometimes successfully) to degrade or insult us; to bring us low; to say, "I am on high while you are down there below." As a result we in a real sense *lose face* when done a moral injury—one reason why easy forgiveness tends to compromise self-esteem. But our moral relations provide for a ritual whereby the wrongdoer can symbolically bring himself low (or raise us up—I am not sure which metaphor best captures the point)—in other words, the humbling ritual of *apology*, the language of which is often that of *begging* for forgiveness. The posture of begging is not very exalted, of course, and thus some symbolic equality—necessary if forgiveness is to proceed consistently with self-respect—is now present.⁹

According to Murphy's suggestion, then, the performance of the humbling ritual addresses not just the harm but the wrongdoing itself, and it does so by restoring the equality of relations that should have obtained and that the wrongdoing violated.

One potential advantage of Murphy's view is that it explains why the body language of apology should exude deference and humility: the performance of subordination (somehow) restores moral equality. As far as Murphy tells us, however, the performance of subordination could be entirely insincere yet still do its normative work. Murphy says that "in the best of cases [apology] is likely to be a way of manifesting repentance."¹⁰ However, he does not explain

9 Murphy, "Forgiveness and Resentment," 28. Note that Murphy's rich account can also be seen as a source of the reassurance view.

10 Murphy, "Forgiveness and Resentment," 28.

how repentance can be manifested through an action with canonical features. Perhaps a proponent of the performing deference view could claim that the subordination of the wrongdoer is genuine only if the wrongdoer has internalized their inferior position and thus feels deferential as well as acting deferentially. Even with this addition, however, the view does not really explain how it is possible for such internalization to take place or how it can be that the wrongdoer authentically expresses the appropriate emotions through those symbolic actions.

Moreover, whether that strategy is plausible depends on the central question of whether the performing deference view is a good explanation of the characteristic normative effects of apology. While it is a strength of Murphy's view that it sees the characteristic role of apology as directly addressing the normative situation of wrongdoing, it is unclear how the fact that the wrongdoer performs a symbolic action would correct that situation. The wrongdoer, according to Murphy, goes through a ritual with a certain form, and let us grant that the form of the ritual is symbolically adequate to the nature of the wrongdoing (though I will dispute this below). Nevertheless, Murphy tells us nothing about why we should believe that such a performance could rationally be taken to bring about normative change.

In order to address this latter criticism, I suggest, we need to see apology not as a symbolic performance to no purpose but rather as a canonical action that brings about distinctive changes in the normative situation.¹¹ On this interpretation, the deferential behavior, along with the other canonical elements of apology, would be the vehicle through which the power to change the normative situation is exercised. The idea here would be of a certain sort of normative transfer: the situation of wrongdoing has brought about an imbalance in the proper distribution of respect, and the normative function of the symbolic performance involved in apology would be to rebalance things by taking away an excess and using it to restore a deficit.¹²

However, even with the performing deference view strengthened in this way, the key to the view remains its diagnosis of the initial moral situation addressed by apology: that wrongdoing brings low a moral subject who, because of basic

11 There are various ways to develop this idea, and in this paper, I do not commit myself to any particular framework. For instance, we might develop it as the idea of an Austinian performative. See Austin, "Performative Utterances." Or we could think of it as a normative power. See Raz, *Practical Reasons and Norms*, ch. 3; Owens, *Shaping the Normative Landscape*; and Bennett, "The Alteration Thesis." The idea that apology is a performative is most explicit in Helmreich's account of apology as "stance-taking" ("The Apologetic Stance"), though see also Martin, "Owning Up and Lowering Down."

12 See Bovens, "Apologies," who also uses this language.

equality, should never have been brought low, and who can be raised up again by the performance of subordination. The performing deference view falls apart if its diagnosis of the initial moral situation is faulty. For in that case, it would have no good explanation of why the performance of symbolically deferential behavior is necessary in the first place and thus no explanation of why that particular normative transfer is called for. As Jean Hampton brings out, however, Murphy's view about the moral situation addressed by apology is highly problematic. It seems to rest on the idea that the victim has in some way *actually been made less than equal* by virtue of the wrongdoing.¹³ Hampton argues that we should reject the idea that people can have their moral status altered by wrongdoing. She argues that we should rather hold the Kantian belief that moral status is unconditional, and that it is an implication of this Kantian belief that the victim cannot really have been lowered in their status. Murphy's view, according to Hampton, would have to rest on something more like a Hobbesian view on which moral status is a limited resource we compete and fight for, and where our ranking in the struggle can go up and down. Since this Hobbesian view is unacceptable, it follows that Murphy's view must be the wrong diagnosis of the normative situation that apology addresses, the wrong account of the normative effects of apology, and the wrong account of the symbolism of apology.

I have now looked at two influential recent accounts of the normative effects of apology. I argued that the reassurance view cannot explain why apology has canonical features, and that in seeking to explain apology as repairing harm rather than addressing wrongs, it cannot capture what is distinctive in apology. While the performing deference view represents an advance because it sees apology as addressing wrongdoing rather than harm, it misidentifies the need for apology and hence misinterprets its symbolism and distinctive normative effects. Furthermore, it provides no account of how symbolic performance can be both expressive and normatively powerful. In attempting to improve on the performing deference view, we need a better understanding of what it is for some action to be expressive of emotion, how expressions of emotion can involve canonical features, how the canonical features of expressions of emotion involve symbolism, and how the canonical features of an expression of emotion can become vehicles for the exercise of powers to alter the normative situation. Once we have a better understanding of these issues, we can then see how this might apply to the case of apology. I will start with an account of *expressive action*.

13 Hampton, "Forgiveness, Resentment and Hatred."

III

According to Jenefer Robinson, the core idea of “expression of emotion” is “a piece of behavior that manifests or reveals that emotion in such a way that we can not only infer from the behavior to the emotion but also perceive the emotion in the behavior.”¹⁴ “Ex-press” is here simply for something within to be pushed out. While I do not want to quibble about core meanings, the idea that I am interested in is somewhat different and closer to what we mean when we say that a piece of art is *expressively powerful*.¹⁵ For Robinson’s core idea of expression does not yet draw a distinction between *symptoms* of emotion and attempts to *give those emotions expression*. It does not distinguish, in other words, between behavior that merely *betrays* our emotions and behavior that is *expressive of* the emotion. There are many acts (as well as nonactions such as blushing and sweating) that may be caused by our emotions but are not expressive of them. Take a case in which I see a dark shape looming toward me as I walk through a darkened alleyway and my fear motivates me to put my hands up in a defensive position. Anyone witnessing this situation will be able to read my fear from my actions. Nevertheless, my behavior is not expressive of my fear: I do not *give* expression to my fear. By contrast, artworks can be attempts to give expression to emotions by the creation of objects or performances with properties that are expressively powerful in relation to some understanding of a situation. Furthermore, what is true of artworks can be extended to actions: when Christians kneel in church, what they do *is* (or can be) expressive of their sense that they are in the presence of a being whose worth is incomparably higher than their own.¹⁶

The point is not simply that the expression of reverence among churchgoers is conventional whereas the expression of fear in the alleyway is not. Rather

14 Robinson, *Deeper Than Reason*, 258.

15 For the account of expressive action developed here, see Bennet, “Expressive Actions,” “The Problem of Expressive Action,” and “How and Why to Express the Emotions.” The sense of “expression” delineated in the text is not identical with the type of expression explained in works in the philosophy of language such as Green, *Self-Expression*; and Davis, *Meaning, Expression and Thought*. More germane to my concerns is a debate in philosophy of action about expressive action, or action out of emotion, initiated by Hursthouse, “A-rational Actions.” While some action out of emotion is not expressive in the sense of being expressively powerful, I argue that there is a wide class of expressively powerful actions and, furthermore, that their expressive power is a good (rational) explanation of why we do them. See also, e.g., Betzler, “Expressive Actions”; and Döring, “Explaining Action by Emotion.”

16 As it happens, I am not a Christian, but I hope readers Christian and non-Christian can understand the point of the example.

the key point is that whereas the defensive posture is simply *caused* by my fear, the expression of reverence is expressively *powerful* in relation to the content of the feelings experienced by the churchgoers (in the way that artworks can be expressively powerful in relation to such content). In other words, the form taken by the expression of reverence (that of kneeling or lowering oneself) can be seen as *reflecting* or *capturing* the content of the relevant attitudes—that is, the perception of the incomparably higher worth or value of the Divine. By contrast, when I assume a defensive posture out of fear, I am simply trying to defend myself. There *could* be acts that are expressive of fear: for instance, if before I give a talk at an important conference, I act as though my legs have gone to jelly. But the point is that not all of those acts that betray, or are caused by emotion, are expressive of that emotion in the sense in which we are interested here. What is distinctive of acts that are expressive of emotion is that they are expressively powerful by virtue of the fact that the form they take reflects important elements of the situation at which they are directed.

If this is correct, then we have established that behavior being caused by emotion is not *sufficient* for that behavior to be expressive of that emotion—not in the sense of being expressively powerful in relation to the content of that emotion. However, neither is being caused by the emotion *necessary* for the act to be expressively powerful. Acting as though my legs have gone to jelly can be a powerful expression of fear even if I am not actually feeling fear at the time I engage in this action. Even if I am feeling perfectly confident about my talk, for instance, I might do the wobbly legs routine in order to indicate that I understand that this is an important event and the kind of thing it is quite appropriate to be nervous or fearful about. Perhaps I do this in order to show solidarity with other speakers who are feeling more nervous than I am. Thus, we have the possibility that an expressive act comes to take on social meaning that does not depend on the motivations of the agent.¹⁷

I now want to argue that an act is expressive of some emotion insofar as it is a powerful *symbol* of that emotion.¹⁸ What I mean by “symbol” here is not that the behavior simply *denotes* fear according to some conventional scheme of reference (as an ox in a painting of a saint may denote Luke, and a winged lion, Mark) but rather, to adopt Nelson Goodman’s distinction between types of reference, that it (metaphorically) *exemplifies* it.¹⁹ To exemplify, for Goodman, is to refer to a property by possessing it, as a sample of cloth refers to the

17 Anderson and Pildes, “Expressive Theories of Law.”

18 Bennett, “Expressive Actions.”

19 Goodman, *Languages of Art*, 85. See also the discussion in Eldridge, *An Introduction to the Philosophy of Art*, ch. 4.

cloth itself. A property is expressive of something, on Goodman's view, when it metaphorically exemplifies it. Thus, I will say that some action (or other vehicle) is expressive of an emotional state when it symbolizes those features of the situation that that particular emotion makes salient. That is, it symbolizes the awe- or fear- or remorse- or joy-worthy features of the situation: those features that call for that emotion. And it symbolizes those features when it is such that we can (metaphorically) see those salient features *in* (or exemplified by) that action. For instance, in the kneeling we can see the Christian's perception of their situation as one in which their significance is dwarfed by the incomparable worth of the Almighty.

What makes something expressive of an emotion is therefore not a causal link that it bears to an emotional state but rather its *expressive properties*. And if we follow Goodman, we will say that the properties of an action are expressive when they bear a relation to the referent that is not conventional but rather a matter of intelligible gestalt—of what we can intelligibly see or construe as a telling metaphor for the referent. On this sense of “expressive,” it is possible to engage in expressive actions without experiencing the emotion at the time. An action can be expressive of an attitude without being used in that instance by an agent to express their own attitude.²⁰ Nevertheless, there is something about the form of the action—for instance, its symbolic properties—that makes it particularly appropriate as an expressive vehicle for that emotion. Furthermore, as we will see below, for an action to be expressive in this sense, it must be the case that although the actual underlying presence of emotion is neither necessary nor sufficient for the act to be expressive, it is nevertheless quite intelligible—indeed normal—for people in the grip of an emotion to give expression to it by engaging in those actions.

IV

Why would one engage in expressive actions? Action that is expressive is not aimed, in the first instance, at altering the material situation or bringing about some further end. But this does not necessarily mean that it is pointless. Rather, if there is some point to engaging in expressive action, it is what we might call a backward-looking one: it lies in *marking* the situation as important in some distinctive way and attempting to *do justice to it*.²¹ The point of expressive actions is simply to *acknowledge* or *recognize* their significance. This phenomenon of

20 Thus, we also have the possibility of nonsincere expressions of emotion—that is, of acts that present themselves either as (falsely) indicating the presence of the emotion or as indicating the agent's sense of the appropriateness of the emotion to that situation.

21 Bennett, “Expressive Actions.”

“doing justice” to one’s situation (through expressive actions) is important in our life because it enables us to isolate that situation from the ongoing rush of “one damn thing after another” and allows our attitudes to that situation to themselves become an object of scrutiny. Having an action that resonates with those attitudes and in which we can see those attitudes represented allows us to dwell on the situation and what is salient in it. Perhaps we could even say that such action is a vehicle through which we can dwell in and with what is salient in the situation for an intense period of reflection. Expressive action therefore allows us to mark certain situations or events as pivotal or as otherwise out of the ordinary—as something to which special attention must be paid—through actions that resonate with our sense of why those situations are salient.

Expressive actions, once one starts to look for them, are common. They are the kind of thing we do when we welcome or take leave, when we mourn or celebrate or commiserate, when we thank, or (perhaps) when we blame. Sometimes the same action, broadly described, can be expressive of quite different emotions, and it is the context and the emotional tone of the performance—or the way in which the action is performed—that make it clear which emotion is being expressed. For instance, embracing a loved one symbolizes and is expressive both of being pleased to see them and of being sorry to see them go. In both cases, the act of holding them close is expressively powerful in relation to the past or future in which they are far away. Thus, embracing can be carried out in various ways to capture regret or delight. Sometimes we get it wrong and feel that the departing embrace was too cursory, or that we brushed the departure off lightly, that the embrace did not at all capture what we were feeling (or feel now that we are away from the immediate emotional pressure of the situation).

Expressing one’s emotions can therefore be a complex and creative affair. On the one hand, this is because of the complexity of the situations we find ourselves in, and the multiple aspects that may call for different emotions, such that the dominant saliences are unclear, at least prior to deliberation. On the other hand, it is because each of us may bring an individual style to our expressiveness. Expressive acts have to fit in with the ways of acting that are characteristic of the agent. It can take a while for a person to find their expressive style, and some people are better at it than others. However, even when the emphasis is on finding our own way of expressing what we feel about a situation, we do so against shared background understanding of paradigm scenarios that (perhaps relative to some community of interpreters) provide canonical understandings of which actions are expressively powerful in relation to which situations.²²

22 De Sousa, “The Rationality of Emotions.”

When Christians enter a church, or a parent says goodbye to the child who is going away, they might look, I have suggested, for some action to do justice to the way they feel. But they will not have to search *too* far. They are not beginning the search *ex nihilo*. Rather, in coherentist fashion, they take up some action that seems right to them on the basis of understandings already there in the culture. Again, this is not simply a question of adopting the local conventions. It is a matter of finding metaphors or symbols compelling and powerful on the basis of the way in which they fit with other aspects of one's inherited background of beliefs and values and vocabulary of other expressive actions. Once we see that expressive actions are grounded in an understanding of the power of certain symbols and metaphors, and that these understandings can be shared and embedded in a culture, there is no difficulty in seeing how it could be that individuals can express their own emotions through canonical actions.

We are now in a position to see how it might be at least possible that an apology could be expressive even though it deploys the canonical actions through which it can generate regular normative effects. The fact that the expressive significance of some acts has become common currency in a culture does not mean that we cannot perform them as authentic expressions of emotion. We can imagine the practice of apology altering and developing if agents came to experience the instituted understandings as inappropriate and unsatisfying, or merely ritualistic. It has to be the case that people can use these socially instituted (or collectively developed) forms for the genuine expression of emotion. A practice might die or be radically altered if an adequate expressive vehicle cannot be found. I take it, however, that, for many of us at least, this is not the case with our practice of apology; it is an interesting feature of our—in many respects highly diverse—society, and the patterns of socialization at work in it, that the very same actions necessary to do the normative work of apology can also be experienced as vehicles for the authentic expression of remorse.

V

It is an interesting question how social and historical development might bring it about that the features of expressive actions could become canonical and under what conditions there might be social pressure toward unanimity. Perhaps people converge on a given set of symbols because of their inherent aptness (given a background of other widely accepted beliefs, values, and symbols); or perhaps the convergence is to be explained by the influence of an established religion that preaches a sacrament of penance; or perhaps there is

something of truth in both of these explanations. However, we might imagine that the conception of expressive power on which a community converges feeds into its conception of ritual observance, delivering a shared sense of the actions that are necessary to mark and do justice to the major normative transitions of human existence. Given this convergence, it may be plausible that, over generations, the form of the ritual widely judged to be appropriate starts to mold our emotions and perceptions, until it is the ritual that comes to form the narrative arc of the emotion, the lack to which it responds and the satisfaction it seeks.

Furthermore, we can perhaps now speculate that it would be plausible that if the expressive action of apology did become canonical, then it could also become intertwined with a community's normative understandings and come to take on the significance of a power through which regular normative effects are brought about. If there are such powers—or if a particular culture develops the idea of such powers—it might come to seem important that the form taken by those actions be not simply arbitrary. After all, the action is an important one that brings about weighty normative changes, and it might seem that the action should reflect that significance. Thus, it might seem that the power could only be exercised by acts that are particularly fitting to the normative situation. Perhaps we could interpret this as the idea that the form taken by those actions should be in some way *continuous* with the normative effects being brought about. (An archaic example of such a symbolically adequate power might be where a courtier bows or kneels to a monarch before approaching them for an audience, assuming a deferential posture in order thereby to make such an approach permissible.)

Furthermore, it might come to seem necessary that one cannot exercise such a power *lightly*, but rather that one has to be in a state of awareness of the significance of the power being exercised. If this were the case, then it would be not only the *form* of one's actions but also the *spirit* in which one does them that has to fit with the gravity of the situation. In such a case, we might say, one changes one's normative situation by means of an action that expressively recognizes the significance of the normative distance to be traversed. If this seems plausible, it suggests that for some powers, we should expect the action that exercises the power not to be one that is arbitrarily specified by convention, but rather one that is expressively adequate. The action has to be expressively adequate both in the sense that its form has to correspond to the normative situation being altered and in the sense that the person who exercises the power has to do so with an awareness of the gravity of the situation. In order to exercise the power, one has to act in the way that someone would who appropriately recognizes the gravity of the situation they are in.

VI

How does this account of expressive actions apply to apology? A full answer to this question would have to explain what kind of expressive action apology is and what characteristic normative effects are brought about by its felicitous performance. It would need to explain how the form taken by the actions required by apology is fitting to the situation of wrongdoing. And it would have to explain how those fitting actions are also fitting to the normative effects apology brings about. Defending such an answer in full would require a further paper. However, it is possible to provide a brief sketch of how an account of apology as an expressive action might be developed.

To start with, consider that a common “protesting” response to wrongdoing consists in a refusal to engage in normal relations with the wrongdoer (perhaps until such time as they put things right). It might be asked why we react to wrongdoing this way, with what P. F. Strawson calls a “partial and temporary withdrawal of goodwill.”²³ Is it simply instinct, or morally arbitrary socially constructed behavior? Or is it in some way fitting to the situation? We can explain the intuition that such behavior is fitting by understanding such distancing as an expressive action in the terms just outlined. The distancing is expressive of emotions of condemnation when it is performed as a compelling way of doing justice to the salient features of the situation of wrongdoing. The distancing symbolizes (metaphorically exemplifies) the normative situation of an agent who has violated a fundamental norm of the moral community to which they belong as a self-governing member. On the basis of such membership, the agent would normally be due certain distinctive marks of respect and recognition. Distancing is expressively powerful because it consists in a partial and temporary withdrawal of that respect and recognition, which is carried out because it marks the fact that the wrongdoer has done something that members of the moral community should be committed not to doing. The wrongdoer is not expelled from the community. Expulsion would be entirely the wrong symbolism. They are still within the community, fully subject to its norms, and deserving of the recognition due to its members. But they have acted in a way that a member of the community should have seen as impermissible. Hence, withdrawing recognition is a way of *reaffirming* the wrongdoer’s membership in the community despite their wrongdoing and, as such, is an apt symbol to capture the moral situation in which the wrongdoing has placed the wrongdoer.²⁴

23 Strawson, “Freedom and Resentment,” 77.

24 For the account sketched here, see Bennett, “The Varieties of Retributive Experience” and *The Apology Ritual*, ch. 5.

What we have talked about so far is someone distancing themselves from *another person* as a way of protesting what they have done. However, if apology is an expressive action, then it must be expressive of an attitude toward *oneself* and one's *own* situation. But we can straightforwardly extend the account just given to explain the canonical features of apology. In the case of remorse or self-blame, the expressively powerful forms of behavior have to do with distancing *from oneself*. One withdraws normal relations, as it were, from oneself. It is this self-withdrawal that accounts for the deferential and humble posture of the remorseful, and the penitential willingness to renounce benefits that would otherwise have been one's due and undertake tasks that would not otherwise have been one's duty. Unlike the performing deference view, which explains the appropriateness of deferential behavior by reference to the need to restore the victim's moral equality, the view I am arguing for is not committed to the claim that the victim has initially been brought low. What is rather at issue is the offender's having committed a wrong that distances them from the moral community.

In engaging in deferential, self-denying behavior, the offender *joins with others* in distancing from the wrongdoer. Thus, as long as it is the wrongdoer's own, authentic remorse being expressed through these symbols of self-withdrawal, apology is expressive not just of distance from their own wrongdoing but also of the fact that the wrongdoer is no longer at odds with the moral community. The wrongdoer has taken up the attitude of condemnation shared by other members of the moral community, seeing their action as to be repudiated, to be dissociated from, and has striven to do justice to that attitude by altering their treatment of themselves accordingly, through penitential actions and postures of humility.

Nevertheless, the fact that apology expresses return to the community is not the end of the matter: the wrongdoer is not back on equal terms, but rather on probation. What apology does is therefore a beginning of the process of return, rather than its end. It is expressive of a commitment to stay within the community. The offender whose postapologetic behavior is not expressive of this probationary status risks undermining the meaning of their apology. Thus, it is typically expressively inappropriate simply to return to a carefree demeanor too quickly after apologizing for a significant wrong.

We can now extend our sketch to show how it might be argued that there is a relation of fit between this symbolism and the normative effects that apology purports to bring about. Apology, it might be suggested, has two characteristic normative effects. First, it alters the normative situation so that others can resume normal relations with the offender without moral compromise. And second, it gives the addressee of the apology special rights that the offender

answer to them for their conduct in relation to such desistance: a kind of special oversight authority over that conduct. These normative effects echo the expressive properties of the action whereby the offender dissociates from the earlier self who performed the wrongful action. The vehicle of apology—the actions necessary for the exercise of this power—are not conventional and arbitrary but rather fitting to the situation in which they are exercised. It is the absence of an act with these normative effects that explains the sense of dissatisfaction articulated in *Missing Apology*.²⁵

The account I have sketched here would explain how apology can be expressive of remorse, how it can be restorative, and why it involves some commitment regarding future behavior. If we want to give a name to this account of apology as an expressive action, I would describe it as a theory of apology as *dissociation*.²⁶ In apology, one performs actions that are expressive of dissociation from one's past wrong and reassociation with the moral community. And when those expressive actions express one's genuine remorse, they could, if the view sketched here is plausible, have the normative effect of dissociating oneself from one's past wrong and returning one to the community, albeit with a probationary status, thereby allowing others to resume normal relations without moral compromise.

VII

In this paper, I have looked at how we might rationalize our practice of treating apologies as bringing about distinctive normative effects. Since apology involves a set of canonical actions but is also expressive of remorse, I argued that it is necessary to understand it as an expressive action. Having explained what expressive actions are, I sketched a view on which the expressively powerful, or expressively adequate, way to do justice to wrongdoing is *dissociation*—that is, not to treat the offender normally, and specifically to withdraw

25 I deny that apology can have its backward-looking normative effect of making uncompromised resumption of relations possible unless it is an expression of remorse. However, perhaps an apology can have its forward-looking effect of transferring oversight authority *without* being given sincerely. That is, it may be plausible to say that even the person who gives an apology grudgingly has, simply by making the apology, thereby created a new normative relation to the addressee in virtue of which they can appropriately be called to account by the addressee for failing to take relevant steps toward desistance.

26 For some further thoughts about dissociation, see Bennett, "Complicity and Normative Control." The notion of dissociation through apology is also appealed to in Hieronymi, "Articulating an Uncompromising Forgiveness."

the respect to which they would otherwise have been entitled. Dissociation, I suggested, is what the offender does to himself when he apologizes.²⁷

University of Sheffield
c.bennett@sheffield.ac.uk

REFERENCES

- Anderson, Elizabeth, and Richard Pildes. "Expressive Theories of Law: A General Restatement." *University of Pennsylvania Law Review* 148, no. 5 (May 2000): 1503–75.
- Austin, J. L. "Performative Utterances." In *Philosophical Papers*, edited by J. O. Urmson and G. Warnock, 233–52. Oxford: Clarendon Press, 1961.
- Barnum-Roberts, Brooke Natalie. "Apologizing without Regret." *Ratio* 24, no. 1 (March 2011): 17–27.
- Bennett, Christopher. "The Alteration Thesis: Forgiveness as a Normative Power." *Philosophy and Public Affairs* 46, no. 2 (Spring 2018): 207–33.
- . *The Apology Ritual: A Philosophical Theory of Punishment*. Cambridge: Cambridge University Press, 2008.
- . "Complicity and Normative Control." *Monist* 104, no. 2 (April 2021): 182–94.
- . "Expressive Actions." In *The Expression of Emotion: Philosophical and Legal Perspectives*, edited by Catherine Abell and Joel Smith, 73–94. Cambridge: Cambridge University Press, 2016.
- . "How and Why to Express the Emotions: A Taxonomy of Emotional Expression with Historical Illustrations." *Metaphilosophy* 52, no. 5 (October 2021): 513–29.
- . "The Problem of Expressive Action." *Philosophy* 96, no. 2 (April 2021): 277–300.
- . "The Varieties of Retributive Experience." *Philosophical Quarterly* 52, no. 207 (April 2002): 145–63.

27 Versions of this paper were presented to audiences at Stockholm, Munich, and Essen, and to the Society of Applied Philosophy Annual Conference in Cardiff. I would like to thank those who gave me useful feedback at those events, in particular: Monika Betzler, Susanne Boshammer, Andreas Brekke Carlsson, Peter Chau, Garrett Cullity, Helen Frowe, André Grahle, Charles Griswold, Oliver Hallich, Richard Healey, Bennett Helm, Sofia Jeppsson, Julian Jonker, Ellie Mason, Ben Mathieson, Leonhard Menges, Per Milam, Jennifer Page, Glenn Pettigrove, Linda Radzik, Stefan Riedener, Peter Schaber, Adam Slavny, and Bill Wringe. I have also benefited from very helpful input from referees.

- Betzler, Monika. "Expressive Actions." *Inquiry* 52, no. 3 (2009): 272–92.
- Bovens, Luc. "Apologies." *Proceedings of the Aristotelian Society* 108, no. 1 (October 2008): 219–39.
- Cohen, Andrew. "Vicarious Apologies as Moral Repair." *Ratio* 30, no. 3 (September 2017): 359–73.
- Davis, Paul. "On Apologies." *Journal of Applied Philosophy* 19, no. 2 (2002): 169–73.
- Davis, Wayne A. *Meaning, Expression and Thought*. Cambridge: Cambridge University Press, 2002.
- de Sousa, Ronald. "The Rationality of Emotions." In *Explaining Emotions*, edited by Amélie Oksenberg Rorty, 127–52. Berkeley: University of California Press, 1980.
- Döring, Sabine A. "Explaining Action by Emotion." *Philosophical Quarterly* 53, no. 211 (April 2003): 214–30.
- Eldridge, Richard. *An Introduction to the Philosophy of Art*. Cambridge: Cambridge University Press, 2003.
- Gill, Katherine. "The Moral Functions of an Apology." *Philosophical Forum* 31, no. 1 (Spring 2000): 11–27.
- Goodman, Nelson. *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis: Hackett, 1976.
- Govier, Trudy, and Wilhelm Verwoerd. "The Promise and Pitfalls of Apology." *Journal of Social Philosophy* 33, no. 1 (Spring 2002): 67–82.
- Green, Mitchell. *Self-Expression*. Oxford: Oxford University Press, 2007.
- Hampton, Jean. "Correcting Harms versus Righting Wrongs: The Goal of Retribution." *UCLA Law Review* 39 (1992): 1659–702.
- . "Forgiveness, Resentment and Hatred." In Murphy and Hampton, *Forgiveness and Mercy*, 35–87.
- Helmreich, Jeffrey S. "The Apologetic Stance." *Philosophy and Public Affairs* 43, no. 2 (Spring 2015): 75–108.
- Hieronymi, Pamela. "Articulating an Uncompromising Forgiveness." *Philosophy and Phenomenological Research* 62, no. 3 (May 2001): 529–55.
- Hursthouse, Rosalind. "Arational Actions." *Journal of Philosophy* 88, no. 2 (February 1991): 57–68.
- Joyce, Richard. "Apologizing." *Public Affairs Quarterly* 13, no. 2 (April 1999): 159–73.
- Lazare, Aaron. *On Apology*. Oxford: Oxford University Press, 2004.
- Martin, Adrienne M. "Owning Up and Lowering Down: The Power of Apology." *Journal of Philosophy* 107, no. 10 (October 2010): 534–53.
- Murphy, Jeffrie G. "Forgiveness and Resentment." In Murphy and Hampton, *Forgiveness and Mercy*, 14–34.

- Murphy, Jeffrie G., and Jean Hampton, *Forgiveness and Mercy*. Cambridge: Cambridge University Press, 1988.
- Owens, David. *Shaping the Normative Landscape*. Oxford: Oxford University Press, 2012.
- Pettigrove, Glen, and Jordan Collins. "Apologizing for Who I Am." *Journal of Applied Philosophy* 28, no. 2 (May 2011): 137–50.
- Raz, Joseph. *Practical Reasons and Norms*. 2nd ed. Oxford: Oxford University Press, 1999.
- Robinson, Jenefer. *Deeper Than Reason: Emotion and Its Role in Literature, Music, and Art*. New York: Oxford University Press, 2005.
- Smith, Nick. "The Categorical Apology." *Journal of Social Philosophy* 36, no. 4 (Winter 2005): 473–96.
- . *I Was Wrong: The Meanings of Apologies*. Cambridge: Cambridge University Press, 2008.
- Strawson, Peter. "Freedom and Resentment." In *Free Will*, edited by Gary Watson, 59–80. Oxford: Oxford University Press, 1982.
- Tavuchis, Nicholas. *Mea Culpa: A Sociology of Apology and Reconciliation*. Stanford, CA: Stanford University Press, 1991.
- Walker, Margaret Urban. *Moral Repair: Reconstructing Moral Relations after Wrongdoing*. Cambridge: Cambridge University Press, 2006.