

The Social Role of Reactive Attitudes

1. Introduction

First highlighted by Peter Strawson (1962), reactive attitudes such as blame, empathy and forgiveness are attitudes held in response to the actions of an agent.¹ Sometimes these reactive attitudes are used as an instrument of power, not different in kind from physical power. For instance, we are familiar with teachers using blame to inspire their students with values that have a lasting effect on their behaviour. Just as with physical coercion, this power can be used for good or ill (Fricker 2016, 181). It is wielded by sports coaches, cult leaders and other teachers, to educate or to indoctrinate. This phenomenon raises difficult social and political questions, but it also poses an important puzzle for the philosophy of blame, and of reactive attitudes more generally.

In this essay, I raise a puzzle for Miranda Fricker's contribution on blame (§2). Fricker understands blame as a persuasive attempt to align the reasons of blamed and blamer (Fricker *ibid.* 175). However, if blame is just alignment of reasons, then it is difficult to draw the intuitive distinction between problematic and unproblematic uses of blame. What is the difference between a schoolteacher's proper use of blame, and a cult leader's improper use, if both just require alignment of values (§3)? What does '(im)proper' mean in this context? My aim is to pin down the intuitive difference while retaining the virtues of Fricker's original account.

Moral realists can solve this puzzle easily. A moral realist – briefly, someone who believes that moral sentences are truth-apt, and true iff they correspond to mind-independent moral facts – can say that blame is used properly only if it attempts to align our reasons with the moral facts (§4). However, I believe that Fricker's position is stronger if she is uncommitted to moral realism. I will show that Fricker can solve the puzzle without appealing to realism by appealing to a metaethically neutral value, namely freedom, instead (§5).

2. Communicative Blame

Blaming is a complex, internally diverse practice associated with multifarious affective, cognitive and behavioural states. Types of blame are unique to particular relationships, such as between friends, from teacher to student and from employer to employee. There are many accounts of blame and blameworthiness, analysing them in terms of judgment, emotion or something else. Different accounts target different elements of blame: the concept of blame, the attitude of blame and the practice of blame. Traditionally accounts of blame propose necessary and sufficient conditions, which are then subjected to possible

¹ For current purposes, an attitude is an intentional mental state, a mental state directed at some content.

counterexamples.² However, as Martha Nussbaum remarks, the diversity of blame undermines the hope for lowest-common-denominator definitions. Any set of necessary and sufficient conditions will probably fail to capture blame's complexity, for in seeking to cover all cases, it loses the detail that made blame interesting in the first place (2016, p.260). Fricker's account is an alternative to the traditional approach that captures this complexity without over-simplifying the practice.

Fricker adopts an approach to blame that aims to accommodate its internal diversity. She treats the concept of 'blame' as *familial*, one that denotes various practices which do not necessarily share a single set of central features. Further, she analyses the family members as variations on a *paradigm* case of blame-practice, where the paradigm is to be understood in terms of its function (Fricker *ibid.* 172).

Before I explore her supposed paradigm case, 'Communicative Blame', I distinguish two components of the account. The first is Fricker's argument that Communicative Blame is the explanatorily basic paradigm of blame. The second is the functional analysis of Communicative Blame. How can one non-arbitrarily choose a particular type of blame as the paradigm? It could be the simplest type, or the historically prior type, from which all other types originate. But I don't rely on Fricker's choice of paradigm; I only rely on her analysis of Communicative Blame, with which I am interested independently of whether it can be used to theorise other kinds of blame.

Communicative Blame is the second-person speech-act that judges another's wrongdoing and expresses negative emotion, whose illocutionary aim is the attempt to inspire remorse in the wrongdoer, and whose perlocutionary aim is the attempt to bring about lasting change. I spend the rest of §3 unpacking this definition.

A speech-act is an act performed by a speaker in making an utterance. John Langshaw Austin distinguishes between three types of speech-act; locutionary, illocutionary and perlocutionary acts. A speech-act is locutionary if it involves just the production of an utterance that 'can be classified by its phonetic, grammatical and lexical characteristics'; it is also 'the performance of an act that can be classified by its content' (Longworth 2012, §2.3). A speech-act is illocutionary if it involves locutionary content as well as 'force' (stating, promising, blaming etc.). Two distinct utterances such as my statement that I will return the book tomorrow, and my promise that I will return the book tomorrow, may perform the same locutionary act (they are similar in content) but distinct illocutionary acts (they are distinct in force). A speech-act is perlocutionary if it is classifiable by its 'consequential effects upon the feelings, thoughts of actions of the audience, or of the speaker, or of other persons' (Austin 1962, p.101). My perlocutionary speech-act in promising to return the book tomorrow, if successful, generates the librarian's expectation that the book will be returned. 'Felicity conditions', as they have become known, are the conditions of a speech-act's success, for example, that the librarian trusts me. If she doesn't, then

² For example, Thomas Scanlon's 2008 account analyses blame and blameworthiness in terms of impaired relationships (p.128). Susan Wolf presents a counterexample; we blame family members but familial blame doesn't imply an impaired relationship (2011, p.334).

my attempt to generate an expectation fails. A second-person speech-act is a speech-act targeted at somebody else whom one directly addresses.

What does it mean to say that blame's illocutionary aim is the attempt to inspire remorse? Fricker is developing a thought attributable to Bernard Williams. Williams' reason internalism is the view that A (an agent) has reason to ϕ (some action) only if A could be motivated to ϕ by sound deliberation from S, where S is A's internal motivational states, including drives, passions, values, aims, projects and commitments. Roughly, Williams develops the notion of 'sound deliberation' into knowledge of the empirical facts and valid inference from them, and S, to actions (Williams 1995, p.36). For Fricker, the attempt to inspire remorse in A for ϕ consists in conveying judgments and negative emotions to A, such as disapproval, with the intention of giving A *retrospective* reasons not to have ϕ -ed. Reasons are retrospective iff A has already ϕ -ed, but were A to have such reasons prior to ϕ -ing, they would be less motivated to ϕ . If ϕ is a token of a ϕ -type action, then a retrospective reason not to have ϕ -ed will give A reason to not undertake ϕ -type actions in the future. If successful, the attempt to inspire remorse results in a lasting change in behaviour, which is the perlocutionary aim of blame. If we assume reason internalism, how can we give someone reason (not) to act in a particular way? How, for instance, can parents give their teenage son reason not to bunk off school?

They could give him a new reason by changing his S, such as through punishment, although this will not necessarily be a rational process, since they won't appeal to objects already in his S, which is plausibly a necessary condition of rationality. Or the parents can make the teenager more aware of the reasons already present in his S, claiming that he wasn't paying attention to his own S, by reminding him that he really does enjoy classwork. Or the parents could argue by sound deliberation from his S to (not) ϕ -ing, by providing him with new empirical information and valid inferences from his S. For example, they could appeal to his desire to graduate from university, and the empirical fact that '*no university is going to accept you unless you go to school!*'.

David Lewis' scoreboard metaphor is useful here. For Lewis, unconscious rules regulate conversation, and the rules themselves are up for grabs; such rules work like a common scoreboard between conversational members. Anytime someone utters or presupposes P, 'P is accepted' is added to the scoreboard unless other conversational members dispute it. Acceptance means different things depending on the type of conversation (the type of 'language game' being played). 'P is accepted' could mean variously that 'P is true', 'P should be taken as true for the time being', or 'P is to be imagined', relative to context (Lewis, 1979).

There are many types of conversation; we are considering those in which I attempt to give someone a reason for action, where acceptance to the scoreboard means 'is part of the normative landscape'. The normative landscape is just a metaphor for our fluctuating obligations and rights owed to and from one another (Owens, 2012, pp. 1-21). The scoreboard is a shared representation of the normative landscape. Speech acts can shape the normative landscape or communicate that the normative landscape has a particular structure. For example, the illocutionary force

in uttering a promise *shapes* the normative landscape; it confers a duty to fulfil that promise. But when we blame by saying ‘how dare you!’, we also *communicate* by stating that the normative landscape has a particular structure, for example that the blamed person has a duty to apologise. A communicative act is successful only if there is uptake, where uptake is to be understood as acceptance in the abovementioned sense. There is uptake of a communicative act only if the felicity conditions of that speech-act are fulfilled. The felicity conditions of blame include empathetic listening, respect and taking the blamer seriously. Without these conditions, the blamed person will not accept the blamer’s blame; it will be misinterpreted, ignored or dismissed (Fricker, *ibid.* 172). But if blame is successful, then the blamed person will accept the blame by accepting a retrospective reason not to have ϕ -ed, such that blamed and blamer’s reasons come into alignment.

Fricker’s account of Communicative Blame is worth defending because it is a functional analysis, which thereby provides a useful method for evaluating any particular person’s use of blame, according to how well they fulfil the aim of blame. The illocutionary aim of Communicative Blame is the attempt to inspire remorse, and the perlocutionary aim is the attempt to generate a lasting change in behaviour; any person’s act of blaming is successful to the extent that it achieves these goals.

However, Fricker’s account is susceptible to a challenge that I will state as a puzzle. If blame just involves the alignment of reasons, then how can we mark the intuitive difference between proper and improper uses of blame? A parent who blames their skiving teenager, and a radicalising cult leader who blames unfaithful disciples, are both bringing reasons into alignment. What, if anything, distinguishes their use of blame?

3. A puzzle about proper and improper blame

Blame aims at aligning reasons and given that we defined reasons in terms of values (among other things), blame is an attempt to align *values*. Blame is a kind of teaching, where the blamer tries to teach values to the blamed until there is alignment. But not all cases of teaching are good, as illustrations will show. Parents blame their skiving teenagers in order to get them to adopt mature values. Teachers communicate with and blame tardy students in order to make them diligent. Feminists blame patriarchal misogynists in order to make them more self-aware. Religious fundamentalists blame open-minded behaviour among their followers in order to encourage the adoption of fanatical values. Cult leaders blame vulnerable disciples to get them to adopt values of loyalty and paranoia. Historically, fascist teachers blamed disobedient children to make them loyal to Nazism.³

Plausibly the former cases are proper and the latter, improper. I don’t mean that nothing can go ‘wrong’, in the broadest possible sense, with the former cases; there is just an intuitive sense in which the latter cases are improper, unlike the former. Both proper and improper cases involve alignment of reasons, so alignment itself cannot tell them apart. Brainwashing is worse than blackmail, which makes the improper cases

³ We can give analogous cases for praise instead of blame, where praise may also be understood in terms of alignment, but blame is my current focus.

particularly tragic. Forcing someone to steal at gunpoint is bad but getting someone to steal by indoctrinating them with a warped morality is worse. It doesn't just violate one's actions; it also invades one's character and sense of self. The possibility of being forced to *want* to do evil repulses me more than the possibility of being forced to do evil.

This puzzle is salient when we consider the controversial post-WW2 're-education' of soldiers at Wilton Park, Buckinghamshire. During this period over 4000 German POWs attended classes on liberal democracy including conversations with Bertrand Russell and Lady Astor, in an attempt to eradicate Nazism (Mayne 2003, p.6). Suppose one of the methods used to re-educate the Nazis involved expressions of blame. Was this use of blame proper, in the relevant sense? What conditions must have been fulfilled for it to have been a proper usage? What, if anything, distinguishes liberal *re-education* from authoritarian *indoctrination*? Recall that I am looking for a necessary condition of proper blame, not a sufficient condition.

This puzzle arises for Ideological Blame as well as Communicative Blame. Ideological Blame is self-directed blame that attempts to internalise structural values. Internalisation of values works much the same way as 'giving reasons' worked for Williams but is first-personal; it includes giving a value to oneself, in the same way that we give another agent reasons (not) to ϕ . A value is structural (or 'systemic') iff it is culturally inherited, its effects are so embedded in social life as to become habitual, it's 'opaque', which is to say that it's hard to recognise, or any one of these.⁴ I assume that there are such values, sometimes called 'cultural norms' or 'ideological values' and that, for better or worse, we sometimes blame ourselves for not fulfilling them (Geuss 1981, pp. 9-10). Ideological Blame is harder to spot than Communicative Blame because the values are opaque.

For example, some 'effeminate' men blame themselves for not meeting the inherited standards of cocksure masculinity, and some naturally-outspoken women blame themselves for not fulfilling the inherited value of meek femininity. There are many structural values and not all of them are obviously bad; we blame our flustered selves for not being 'cool', or for not contributing enough to our community, thanks to the cultural value of 'social responsibility'. There are a diversity of forces operating for good and ill, so the same puzzle arises for Ideological Blame. What distinguishes the proper cases (thoughtfully blaming oneself for being socially irresponsible) from the improper cases (blaming oneself for being an outspoken woman)? I assume such an intuitive difference exists. If both cases just involve alignment between one's S and structural values, then alignment alone doesn't differentiate them.

4. The first solution: moral facts

Moral realists answer this question with ease; alignment of reasons is proper only when alignment converges on the moral facts, and sinister when it doesn't. Moral realists claim that moral sentences are truth-apt, and sometimes true, and true iff they correspond to mind-independent moral facts. Moral facts can be used to distinguish between proper and

⁴ This analysis of 'structural' phenomena owes to John Filling's talk at the Moral Sciences Club, Cambridge, October 2017.

improper uses of blame. One might think that liberal teachers, parents and feminists use blame properly, because they speak the truth about moral facts whereas cult leaders, religious fundamentalists and fascists use blame improperly because they speak only falsehoods.

The realist approach needs slight refinement, given an important counterexample. An environmentalist blames amoral corporations for global warming but uses misleading rhetoric and fabricated statistics to aid her project. She's preaching a truth ('companies should emit less carbon'), but she's doing it in the wrong way. Even if she succeeded in changing corporations' behaviour her use of blame would be improper.

The realist should qualify their account, saying that blame is used properly only if it *aims* at the truth. The difference between education and indoctrination is not just whether the subject-matter is true, but also whether the blaming practice aims at truth. The global warming advocate uses blame improperly because she uses misleading rhetoric and so disregards the truth (even though climate change is a genuine, important issue).

Fricker doesn't defend a specific metaethical position in her paper, but she should reject the realist solution for three reasons. Firstly, realism is a substantive, controversial position so her position is more persuasive to the extent that she remains metaethically neutral. Secondly, her theory is more elegant to the extent that it has fewer ontological posits, including the existence of moral facts. Thirdly, (and not conclusively), to the extent that her account of blame fits most naturally with an internalist – and therefore anti-realist – account of reasons, an appeal to moral facts is not in keeping with the 'spirit' of her account. Reason internalism suggests anti-realism because it denies that A could have a reason to ϕ if A doesn't have an appropriate item in their S. It therefore makes A's reasons for action dependent on A's values and psychology; it doesn't appeal to mind-independent moral facts about how we should behave. Consequently, while the realist can solve the puzzle with relative ease, it comes with costs, and a defender of Fricker should first look to see if there are other metaethically neutral solutions.

5. The second solution: freedom

There is a metaethically neutral solution. Blame is used properly only if it preserves a particular kind of *freedom*. In this section, I will outline and defend this alternative solution.

Gerald MacCallum Jr famously claimed that you can express any conception of freedom by specifying the three variables in his 'triadic formula': X is free from Y to Z (MacCallum 1967, 314). In this case, X is any agent that can receive blame, Y is irrational adoption of reasons and Z is rational adoption of reasons. Blame is used properly only if it preserves the blamed's freedom to form their reasons rationally. I can think of at least two ways in which adopting reasons can be irrational (although there may be more); you may adopt reasons not to ϕ as a response to punishment that conditions your S to avoid ϕ -type behaviour, or you may adopt reasons unconsciously by blaming yourself for falling short of opaque values.

Rationality (thus defined) does not analytically include the notion of moral facts, so this solution is metaethically neutral. We could further define rationality in terms of truth, e.g. when A adopts a reason to ϕ , A's adoption of reasons is rational only if 'A should ϕ ' is true. The moral realist can define moral truth in terms of moral facts, but truth needn't be defined this way. Anti-realists can define moral truth as coherence with the rest of an agent's S. This still invokes moral truth, so won't appeal to noncognitivists, who deny that moral sentences are truth-apt. But an appeal to truth appeals to realists and anti-realists alike, since some anti-realists permit the truth-aptness of moral sentences without affirming the existence mind-independent moral facts, such as constructivists (Street 2010, 363-84). Since freedom is a metaethically neutral value, this solution has broader appeal, is more elegant, and remains true to the 'spirit' of reason internalism, unlike the defence that appealed to moral facts.

Freedom deals with the puzzle for Communicative Blame. When a feminist blames a misogynist in the hope that he will adopt egalitarian behaviours, she doesn't seek to persuade him through punishment. Conversely, a cult leader improperly blames his unfaithful disciples if he disrespects their rationality by attempting to change their reasons through physical and social punishment.

Truth is not required for proper blame; only the freedom to be rational is required, and freedom can be present without truth. For example, suppose that a particular religion, that of the ancient Roman pantheon, is false, and that there is a family that still worships the Roman gods. The parents in this family blame their teenage daughter for disloyalty to the faith. One parent expresses blame towards the teenager by threatening her with punishment if she doesn't believe in Minerva. The other parent also blames the teenager for not believing in Minerva, but instead of making threats, gives her books about Roman religion and other religions. The former parent violates the teenager's freedom of rationality and the latter preserves it. On my analysis, the latter parent's use of blame is more proper than the first, even though both are aimed at inculcating reasons that would motivate loyalty to a false doctrine.

Freedom also solves the puzzle for Ideological Blame. We blame ourselves according to culturally-inherited values while preserving our freedom only if we are conscious of such values and conclude that they are worth perpetuating. When we blame ourselves for not being sufficiently socially responsible, our blame is proper only if we have noticed, reflected on and accepted the value of social responsibility. When we blame ourselves for violating traditional gender roles, our blame is improper if those values remain opaque to us. A consequence of this view is that if a woman thinks, upon reflection, that a gender role is appropriate, her self-blame for falling short of the gender ideal is free and therefore proper in the relevant sense.

However sometimes we *need* blame that disrespects freedom. Recalling William's original doctrine, consider cases where a court judge blames a murderous psychopath, or when parents blame very young children for disobedience. We can't persuade them by appealing to dormant reasons in their S, because their S is too minimal and deprived. Instead we directly coerce them whether through court punishments or the naughty step, in the hope that this will change their S. On Fricker's analysis of

Communicative Blame, both the naughty step and court punishments are uses of blame, since blame consists (in part) in the attempt to give another person reasons not to have ϕ -ed, and in both cases our aim is to give infants and psychopaths reasons-through-punishment not to have violated rules. Given that we defined irrationality in terms of reasons-through-punishment, this use of blame disrespects their rationality and therefore their freedom, but is nevertheless required for law-abiding society to be possible in the first place. While blame should be aimed at maximising freedom, total freedom is incompatible with society.

What of the controversial re-education of Nazi soldiers at Wilton Park? On this account, such re-education is proper only if it attempts to change the soldiers' S through rational appeals. If the re-education included expressions of blame that appealed to fabricated statistics and threats of punishment for remaining loyal to Nazism, it would count as indoctrination. Another consequence of this view is that grounding teenagers is improper because it violates their freedom. Parents should appeal instead to their skiving teenagers' pre-existent S, rather than changing it through punishment. Grounding teenagers is proper only if it is necessary in the above sense, the sense in which socialisation of psychopaths and very young children is necessary for society to exist at all.

To summarise, Fricker's account defines blame in terms of its function, which is to align practical reasons. A puzzle for the alignment account is that it can't differentiate, by itself, between proper and improper uses of blame. This is a puzzle for reactive attitudes generally, although blame has been my focus. I have shown that Fricker can solve the problem by appealing to freedom, which is attractive because it doesn't appeal to moral facts. Blame is acceptable only if it respects the freedom to rationally shape one's own reasons.

Bibliography

- Austin, John Langshaw (1975), *How To Do Things With Words*, 2nd edition, ed. James O. Urmson, (Oxford University Press).
- Filling, John (2017), 'On Structural Injustice', a talk presented at the Moral Sciences Club, Cambridge on 24th October 2017.
- Fricker, Miranda (2016) 'What's The Point of Blame? A Paradigm Based Explanation', *Noûs*, 50(1), 165–183.
- Geuss, Raymond (1981), *The Idea of a Critical Theory: Habermas and the Frankfurt School*, (Cambridge University Press).
- Lewis, David (1979), 'Scorekeeping in a Language Game', *Journal of Philosophical Logic*, 8, 339-359.
- Longworth, Guy (2017, 'John Langshaw Austin', *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), ed. E. N. Zalta, URL = <https://plato.stanford.edu/archives/spr2017/entries/austin-jl/>.
- MacCallum, Gerald C (1967), 'Negative and Positive Freedom', *The Philosophical Review*, 76(3), 312-334.
- Mayne, Richard (2003), *In Victory, Magnanimity, in Peace, Goodwill: A History of Wilton Park*, (Frank Cass Publishers).
- Nussbaum, Martha (2016), *Anger and Forgiveness* (Oxford University Press).
- Owens, David (2012), *Shaping the Normative Landscape* (Oxford University Press).
- Scanlon, Thomas (2008), *Moral Dimensions: Permissibility, Meaning, Blame*, (Harvard University Press).
- Strawson, Peter (1962), 'Freedom and Resentment', in *Proceedings of the British Academy*, 48, (Oxford University Press), pp. 1-25.
- Street, Sharon (2010), 'What is Constructivism in Ethics and Metaethics?', *Philosophy Compass*, 5(5), 363-84.
- Williams, Bernard (1995), 'Internal Reasons and the Obscurity of Blame', in his *Making Sense of Humanity and Other Philosophical Papers 1982-1993*, (Cambridge University Press), pp. 35-45.
- Wolf, Susan (2011), 'Blame, Italian Style', *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, ed. R. Wallace & R. Kumar & S. Freeman, (Oxford University Press), pp. 332-346.