GROUNDING RESPONSIBILITY IN APPROPRIATE BLAME

Leonhard Menges

ABSTRACT

When confronted with the question of why it is appropriate to morally blame a person for some bad action, it may seem plausible to reply that she is morally responsible for it. Some authors, inspired by Peter Strawson’s “Freedom and Resentment,” argue, however, that thinking this way is backwards. They believe that a person is morally responsible for some bad action because it would be appropriate to blame her for it. The aims of this paper are to present this account, to highlight some of its important but often overlooked features, and to defend it against pressing objections.

1. INTRODUCTION

Suppose that you are responsible for a moral wrongdoing, say deliberately stepping on your neighbor’s foot. You know that stepping on her foot would hurt her, that she could legitimately demand that you not do it, and that doing it would be wrong. Even so, you deliberately and freely step on her foot. Other things being equal, it seems appropriate for your neighbor to blame you for stepping on her foot. The question I will discuss in this paper is this: What is it for us to be morally responsible for a morally questionable action such as your stepping on your neighbor’s foot?

Inspired by Peter Strawson’s “Freedom and Resentment” (1962/2003), many authors believe that a person’s being morally responsible needs to be accounted for in terms of holding her responsible. And holding people responsible is then taken to be essentially connected with having certain blame emotions toward them, such as resentment, indignation, and, in the case of holding oneself responsible, guilt. R. Jay Wallace, a prominent defender of this line, argues for the thesis that a person s “is morally responsible (for action x) if and only if it would be appropriate to hold s morally responsible (for action x)” (Wallace 1994, p. 91). In what follows, I will refer to this view as Weak Strawsonianism (for a view of this kind see, e.g., Watson 1996/2004; Fischer and Ravizza 1998; Darwall 2006; Schulte 2013).

Weak Strawsonianism is weak in the sense that it is a thesis about necessarily co-occurring facts. It does not say that one side of the biconditional is more fundamental than the other. Inspired by Strawson, some authors seem to accept a stronger claim about metaphysical fundamentality. Wallace is, once again, an important example. He argues that a person’s being morally responsible depends on and is fixed by the appropriateness of holding her responsible (see Wallace 1994, pp. 89, 91, 93). This is the central idea of what I will call Strong Strawsonianism. The aims of this paper are to highlight some of Strong Strawsonianism’s important but often overlooked features (section 2) and to defend it against important objections (sections 3–6).
Before discussing Strong Strawsonianism in detail, let me make two brief remarks. First, I will follow most authors in focusing on responsibility for unexcused, morally questionable actions. Thus I will not discuss moral responsibility for morally neutral or positive actions. Second, I will not discuss the nature of holding responsible. I will assume—following Wallace and others—that holding a person responsible for a morally questionable action is to have one of the blame emotions toward her or to judge that having one would be appropriate.

2. INTRODUCING STRONG STRAWSONIANISM

As the aim of Strong Strawsonianism is to account for what it is to be morally responsible for questionable actions, I will begin with a quick pre-theoretical sketch of what it is that Strong Strawsonianism aims to illuminate. When we excuse a person for an objectionable action, we try to show that she is not morally responsible for it. In everyday life, we say things like “you did not intend to step on your neighbor’s foot,” “you did not want to do it,” “you did not know that you would do it,” “you could not have known that you would do it,” and so on. When we say these things, we are not trying to show that it was not you who stepped on her foot, or that it was not you who caused her pain. Thus we do not deny that you are causally responsible for what happened to her. Instead, what we are trying to show is that you are not morally responsible for stepping on her foot and causing her pain. A certain kind of control and a certain kind of knowledge seem to be necessary for being morally responsible for objectionable actions. Similarly, when we tell someone that the person who stepped on her foot is a young child, we are trying to make her see that no one is morally responsible for what happened to her. Very young children seem to lack the abilities that are necessary for being morally responsible for objectionable actions.

Moreover, our forgiving certain people presupposes that we believe that they are morally responsible for the action we forgive them for. This is one of the important differences between forgetting, suppressing, and re-evaluating on the one hand and forgiving on the other. When we forgive a person for a certain objectionable action, we do not stop thinking that she was morally responsible for it.

Given this sketch of moral responsibility for objectionable actions, let me take a closer look at Wallace’s account in order to introduce Strong Strawsonianism. He explicitly argues for what I call Weak Strawsonianism: a person “is morally responsible (for action x) if and only if it would be appropriate to hold s morally responsible (for action x)” (Wallace 1994, p. 91). Even though theses about necessary and sufficient conditions do not say anything about metaphysical dependency, Wallace claims that the left side of the biconditional depends on the right: he contends that “we must interpret the relevant [responsible] facts as somehow dependent on our practices of holding people responsible” (Wallace 1994, p. 89; emphasis added; see also p. 91). It is important that he takes this dependency to be a metaphysical relation, that is, a relation between the responsibility facts and certain facts about the practice of holding responsible.

The claim that fact x depends on fact y is compatible with the claim that fact y depends on fact x. If both are true, then there is an interdependence relation between x and y. But Wallace does not seem to believe that facts about being responsible and facts about holding responsible are interdependent. He claims that moral responsibility “facts are fixed by the answer to the question of when it is appropriate to hold people responsible” (Wallace 1994, p. 93; emphasis added). The
notion of fixing suggests that facts about responsibility depend on those things that fix them, but these things do not depend on facts about responsibility. Thus the notion of fixing suggests an asymmetry.

In the sentence just quoted, Wallace claims that the responsibility facts are fixed by “the answer” to the question of when it is appropriate to hold people responsible. But he surely does not mean that whether or not a responsibility fact obtains depends on what we actually answer to that question, because our actual answers can be false. It is more plausible to interpret him as suggesting that the responsibility facts are fixed by the normative facts that make a certain answer to that question true. Thus Wallace can be interpreted as claiming that responsibility facts are fixed by facts about when it is appropriate to hold people responsible.

Claims about nothing more than necessary and sufficient conditions—such as Weak Strawsonianism—do not describe any metaphysical dependency between the two sides of the biconditional. But, as I have just shown, Wallace can be plausibly interpreted as arguing for a metaphysical, asymmetrical, dependence relation between the appropriateness of holding a person responsible and her being responsible. Claims to the effect that there is a metaphysical, asymmetrical dependence relation between the two sides of the biconditional are often understood as claims about the grounding relation between the facts:

This is Strong Strawsonianism’s central thesis about moral responsibility. Even though this claim is inspired by Wallace’s account of responsibility, it is not entirely clear that it is the view he accepts. Some passages can also be read as suggesting that a person’s being responsible for some action is identical with it being appropriate to hold her responsible for it. And this claim is incompatible with Strong Strawsonianism because the grounding relation is different from the identity relation. The identity thesis is an interesting alternative to Strong Strawsonianism and to...
the position that responsibility is more funda-
mental than the appropriateness of holding
responsible. However, in what follows, I will
leave this view, as well as exegetical ques-
tions about Wallace’s account aside. Instead, I
will defend Strong Strawsonianism itself against
four important objections.

3. THE NATURALLY FALSE OBJECTION

What I call the Naturally False Objection
starts with the claim that it is “natural to think
that responsibility facts are fixed by features
of the agent and the agent’s action” (Vargas
2004, p. 225). It is, for example, natural to
think that what makes a person morally re-
sponsible for a morally questionable action
is that she really performed that action, that
she knew what she was doing, and that she
had some kind of control over her thoughts
and behavior. The objection then says that
Strong Strawsonianism argues that responsi-
bility facts are fixed by facts about responses
toward the agent, and not by facts about the
agent herself and her action. Thus, one might
conclude, it is natural to think that Strong
Strawsonianism is false.

But once Strong Strawsonianism is under-
stood in the way I developed it in section 2, it
becomes clear that this objection is based on
a misunderstanding. Strong Strawsonianism
says that the fact that it would be appropriate
to hold a person morally responsible makes
it the case that she is responsible. But this
does not rule out the possibility that the fact
that it would be appropriate to hold that per-
son responsible is grounded in further facts,
including facts about the person and her
actions. The only claim about the grounds
of the appropriateness of holding a person
responsible is that Strong Strawsonianism rules
out is that the responsibility of the person (is
part of what) makes it the case that it would
be appropriate to hold her responsible.

Strong Strawsonianism is compatible, for
example, with the claim that what makes it
appropriate to hold a person responsible for
some wrongdoing is that she really performed
it, that she knew what she was doing, that
she had the relevant sort of control over her
thoughts and actions, and that it is appropriate
to blame those agents who have these prop-
erties. Then the fact that it is appropriate to hold
that person responsible for that wrongdoing
makes it the case that she is morally respon-
sible for it. This idea can be illustrated in the
following way, in which each box is a fact and
the arrow stands for the grounding relation:

\[ F \rightarrow \text{It would be appropriate to hold } s \text{ responsible for } x \rightarrow S \text{ is responsible for } x \]

Strong Strawsonianism only consists in the
claim that the fact that the right box stands
for is grounded in the fact that the middle
box stands for. And this claim does not rule
out that the fact that the middle box stands
for obtains because of further facts \( f \). It is
not only possible, but also very plausible
that certain facts about \( s \) and \( x \) will be part of
those facts that make it appropriate to hold \( s \)
responsible for \( x \). And these facts about the
agent and her action therefore play a crucial
role in the explanation of the agent’s being
morally responsible for that action. Moreover,
if one assumes that the grounding relation
is transitive, as many authors do, the Strong
Strawsonian can even claim that \( f \) itself makes
it the case that \( s \) is responsible for \( x \).

The opponents of Strong Strawsonianism
will probably not be satisfied by this reply.
They might claim that Strong Strawso-
nianism is still implausible because “in our
common-sense moral ontology, the property
of responsibility is not dependent on some
further and more basic normative property”

How can we find out what our common-
sense moral ontology involves? One plausible
The approach is to take a closer look at our everyday moral practices and to ask whether participating in these practices commits us to certain ontological claims. If it turns out that we are committed to claims that are in conflict with Strong Strawsonianism, this would be a prima facie reason against this account. Let me focus on our practice of excusing agents.

Strong Strawsonianism would conflict with the practice of excusing agents if claims like “she was not responsible for what she did and therefore it is inappropriate to blame her for it” were essential for that practice. In fact, our practice of excusing is neutral with regard to claims about the grounding relation between moral responsibility and the appropriateness of holding responsible. As I suggested above, we excuse people by saying things like “she did not intend to do it” or “she could not have known,” and these excuses fit nicely with Strong Strawsonianism. According to that view, such excuses are attempts to show that some of the facts that are essential parts of the explanation of moral responsibility do not obtain. Facts about what a person intends to do and about what she knows about her action ground, so the Strong Strawsonian argues, whether or not it would be appropriate to blame her for that action. And if certain facts about what she intends and knows do not obtain such that it would be inappropriate to blame that person, then this explains why she is not morally responsible for that action. In this way, many of our everyday excuses are attempts to explain why a person is not morally responsible for a questionable action. Thus Strong Strawsonianism is not in conflict with our everyday practice of excusing people.

More generally, I do not see why our common sense should be in conflict with Strong Strawsonianism. Perhaps some people have that impression because they confuse claims about the grounds of responsibility with claims about the necessary conditions of responsibility. Strong Strawsonianism implies that being morally responsible is not among the grounds of being an appropriate object of blame. Some might think that it also implies that being morally responsible is not a necessary condition of being an appropriate object of blame. And the latter claim is clearly counter-intuitive. However, Strong Strawsonianism explains why being morally responsible is in fact a necessary condition of being an appropriate object of blame: it says that its being appropriate to blame a person makes it the case that she is responsible. Therefore, it is, according to that view, impossible for it to be appropriate to blame a person for an action that she is not responsible for. To sum up this reply, I see no reason why our common sense should include the claim that moral responsibility (partly) grounds appropriate blame. And this is the only claim that would be in conflict with Strong Strawsonianism.

An important upshot of this discussion is that what distinguishes Strong Strawsonianism from other accounts of the nature of responsibility is not a claim about which facts are metaphysically most fundamental. The distinctive feature of Strong Strawsonianism is a claim about the immediate ground of moral responsibility. It says that the appropriateness of certain responses is the immediate ground of the property of being morally responsible. And this is compatible with the claim that the appropriateness of these responses has other grounds, including facts about the agent and her action.

4. The Not-Interesting Objection

Once it becomes clear that Strong Strawsonianism is not a claim about the ultimate grounds of responsibility, it may be argued that Strong Strawsonianism is not interesting. For, it might be said, what we really want to know when we discuss the nature of moral responsibility is not a claim about which facts are metaphysically most fundamental. The distinctive feature of Strong Strawsonianism is a claim about the immediate ground of moral responsibility. It says that the appropriateness of certain responses is the immediate ground of the property of being morally responsible. And this is compatible with the claim that the appropriateness of these responses has other grounds, including facts about the agent and her action.

More generally, I do not see why our common sense should be in conflict with Strong Strawsonianism. Perhaps some people have that impression because they confuse claims about the grounds of responsibility with claims about the necessary conditions of responsibility. Strong Strawsonianism implies that being morally responsible is not among the grounds of being an appropriate object of blame. Some might think that it also implies that being morally responsible is not a necessary condition of being an appropriate object of blame. And the latter claim is clearly counter-intuitive. However, Strong Strawsonianism explains why being morally responsible is in fact a necessary condition of being an appropriate object of blame: it says that its being appropriate to blame a person makes it the case that she is responsible. Therefore, it is, according to that view, impossible for it to be appropriate to blame a person for an action that she is not responsible for. To sum up this reply, I see no reason why our common sense should include the claim that moral responsibility (partly) grounds appropriate blame. And this is the only claim that would be in conflict with Strong Strawsonianism.
interested in other questions is that most current accounts of the nature of responsibility should not be understood as aiming to identify the ultimate grounds of moral responsibility. Take the claim that a person is responsible for some action if and only if and because she had a certain kind of control \( c \) over it (such as being moderately reason-responsive) and she had knowledge \( k \) (such as knowing all relevant normative and non-normative facts about the action and its consequences). This would only be an account of the ultimate grounds of moral responsibility if \( c \) and \( k \) are not grounded in other facts. But it is quite plausible that one’s having a certain kind of control over one’s action and one’s knowing something are psychological facts at least in part. And it is also quite plausible that psychological facts are at least partly grounded in other facts, such as facts about one’s brain. Assuming this, claiming that a person is responsible because of her \( c \) and \( k \) is not to suggest that \( c \) and \( k \) are the ultimate grounds of her moral responsibility. But it is, surely, very interesting whether or not one’s responsibility really is grounded in \( c \) and \( k \). Thus what we are interested in when we discuss the nature of moral responsibility is not only its ultimate grounds. It is interesting to identify the grounds of responsibility at very different levels.

One might agree that we are not only interested in the ultimate grounds of responsibility and still contend that the level Strong Strawsonianism is focusing on, that is, the immediate ground of moral responsibility, is not interesting. Because what we are really interested in when we discuss the nature of moral responsibility, one might claim, is the compatibility question of whether we can be morally responsible in a deterministic world. And Strong Strawsonianism does not, the objection goes, help to answer that question.

It is true that Strong Strawsonianism is neutral with regard to the compatibility question. But it is not true that we are only interested in the grounds of moral responsibility insofar as they are directly related to the compatibility debate. The question “who is responsible for what?” is important for our everyday lives. A better understanding of the nature of moral responsibility would help to answer that question or, at least, to understand what we want to know when we ask it. Therefore, the nature of moral responsibility is an interesting topic independent of concerns about the consequences of determinism. Moreover, if there is a good account of responsibility that compatibilists and incompatibilists can agree on, then this may also help to frame the discussion between these two camps. Thus the neutrality of Strong Strawsonianism can even be seen as a theoretical advantage (see, e.g., McKenna 2012, for an account of the nature of responsibility that is intended to be neutral with regard to the compatibility question).

Finally, Strong Strawsonianism implies an interesting partial answer to an important question: What are the non-normative facts that make it the case that it is appropriate to hold an agent responsible for some bad action? Strong Strawsonianism narrows down the possible replies in a remarkable way. It says that whatever the facts are that make it the case that an agent is an appropriate target of blame for some action, the fact that she is morally responsible for it is not among them.

To sum up, Strong Strawsonianism is an interesting account of the nature of responsibility. It does not offer an ultimate explanation of responsibility, and it is neutral with regard to the compatibility question. But it is a theory of an important part of the metaphysical structure that constitutes moral responsibility. It would be interesting to know whether it is true.

5. The Reductio Objection

The Reductio Objection focuses on the notion of appropriateness. Usually this objection is presented against Weak Strawsonianism, but if the objection is successful, then it also threatens the stronger version I am defending here.
Wallace claims that the appropriateness at the heart of Strawsonianism should be understood as a moral one. More precisely, he argues that a person’s being responsible depends on it being fair to hold her responsible. But critics, such as Angela Smith (2007), argue that whether or not it is fair to hold a person responsible does not only depend on features that seem intuitively relevant for that person’s responsibility. For example, it seems unfair for a person to hold another person responsible for some wrongdoing if the latter already seriously regrets what she did. Or it seems unfair for a person to blame another for a crime if the former has done exactly the same thing. Now the moral version of Strong Strawsonianism seems to say that its not being fair to hold a person responsible makes it the case that she is not responsible. But this is just absurd in these scenarios. The agents may very well be responsible for what they did even though it is not fair for certain people to hold them responsible.

There have been some promising attempts to reply to the Reductio Objection that I will not discuss here (see Maher 2010; McKenna 2012, chap. 2). Instead, I will sketch two more ways in which the proponent of Strong Strawsonianism could reply. First, the version of Strong Strawsonianism I presented in section 2 does not need to insist that the appropriateness at issue is a moral one. Alternatively, the Strong Strawsonian could argue that the appropriateness should be understood in non-moral terms. Many theories of emotions say that emotions have a representational dimension. Part of what it is to, say, fear something is to represent that something as dangerous. The representational dimension of an emotion makes room for a non-moral, representational evaluation of an emotion (see, e.g., D’Arms and Jacobson 2000; Graham 2014). Being afraid of a dog is correct or fitting if the dog is dangerous and it is incorrect or unfitting if the dog is not dangerous. Thus emotions, including the blame emotions, can be appropriate in the sense of being representationally correct, and they can be inappropriate in the sense of being representationally incorrect. Now, the Strong Strawsonian could argue that a person is morally responsible for something if and only if and because it would be correct to have the blame emotions toward her. Then one could describe the cases sketched above in the following way: even though it may be unfair for a certain person to blame someone who regrets what she did, it may nonetheless be representationally correct to have one of the blame emotions toward her. As this grounds responsibility, it follows that the agent really is morally responsible for what she did.

Second, the interpretation of appropriateness in terms of correctness can also be the basis for defending the view that the appropriateness at the heart of Strong Strawsonianism is a moral one. The first step toward such a defense would be to distinguish between different notions or kinds of desert. Then one would have to identify the kind of desert that is necessarily connected with the account of correctness or fittingness I just sketched. According to this sense of desert, a person deserves to be blamed if and only if it would be correct to have the blame emotions toward her (see, e.g., King 2012, 2014; Clarke 2013). Finally, one would have to show that this notion of desert is a moral notion, just like the notion of fairness. It seems quite plausible that there is such a notion of desert, and this notion seems to be in play when we say that a person deserves gratitude for helping us, compassion for her suffering, or honor for her success. If having one of these attitudes toward a person is fitting, then the person seems to deserve (in a certain sense) to be the object of these attitudes.

If there is such a notion of desert, then the Strong Strawsonian can say that it may be unfair for a person to blame a wrongdoer who seriously regrets what she did, but she nonetheless deserves (in that sense) to be blamed for it. And the Strong Strawsonian could add
that a person’s deserving (in that sense) to be blamed for an action makes it the case that she is morally responsible for it.

This is not the place to analyze these notions of desert and correctness or fittingness in detail. But I have shown that there are plausible ways to reply to the Reductio-Objection.

6. Why Strong Strawsonianism?

I have presented Strong Strawsonianism, and I have defended it against three pressing objections. But are there positive reasons to accept it? The opponent might still claim that it is natural to believe that being morally responsible is metaphysically more fundamental than being appropriately held responsible. And she might contend that the Strong Strawsonian bears the burden of proof: as long as there is no positive reason to adopt that view, she could say, we can simply leave it aside as a—curiously enough—not-yet-refuted but somehow counterintuitive account of moral responsibility. What can the Strong Strawsonian reply to that final objection?

First, the Strong Strawsonian should reply that the preceding discussion suggests that claiming that responsibility grounds appropriate blame is not as natural as her opponent seems to think. There is no obvious positive reason for accepting this account, and there is an interesting and defensible alternative to it. This is, admittedly, not a positive reason for Strong Strawsonianism, but it undermines the apparent certainty that responsibility grounds appropriate blame.

Second, Strong Strawsonianism elegantly captures and illuminates an idea that many authors find attractive. This is the idea that responsibility comes into the world because of some facts about the moral community and its practice. An agent is not only morally responsible, the thought goes, because she has certain mental properties; she also has to be integrated in a moral community and its practice in a specific way.

Now, the opponent of Strong Strawsonianism might simply reject that a theory of moral responsibility should be able to account for that idea. And this is not the place to argue that an account of moral responsibility that cannot make sense of it is, therefore, inaccurate. However, many authors find that picture compelling (see, e.g., Strawson 1962/2003; Watson 1987/2004; Darwall 2006; McKenna 2012). And for them, Strong Strawsonianism is a natural option. Take Wallace again. He believes that the opponent account that responsibility facts are more fundamental than facts about appropriately holding responsible is committed to there being a “realm of moral responsibility facts, inhering in the fabric of the world completely independently of our activities and interests in holding people responsible” (Wallace 1994, p. 88). However, he believes that there are no such facts: “My worry . . . is that I cannot see how to make sense of the idea of a prior and independent realm of moral responsibility facts” (Wallace 1994, p. 88). His alternative is that an agent’s being morally responsible for some action is fixed by facts about her being part of a moral community and by facts about the community’s practice.

Strong Strawsonianism spells out this idea in a plausible way. It does not say that facts about our actual practice make it the case that an agent is responsible because our actual practice can be misguided. It says that whether or not an agent is responsible is grounded in normative facts that would ideally guide our practice. An agent is responsible for some bad action, the account says, because it would be appropriate for others to interrelate with her in a certain way, namely by holding her morally responsible for that action. Thus what speaks in favor of Strong Strawsonianism is that it captures and illuminates an idea about the nature of moral responsibility that many find intuitively attractive.

University of Lübeck
I presented earlier versions of this paper at the workshop Moral Responsibility, Amsterdam, October 2014, and at the Colloquium for Practical Philosophy and Ethics at Humboldt-Universität zu Berlin, December 2013. I thank the participants for helpful discussions. I am particularly grateful to Hannah Altehenger, Sarah Buss, Roger Crisp, Claire Davis, Simon Gaus, Jan Gertken, Benjamin Kiesewetter, Andreas Müller, Peter Schulte, Thomas Schmidt, Jay Wallace, and two anonymous referees for APQ, for comments and discussions. Work on this paper has been supported by the Studienstiftung des deutschen Volkes.

1. In this paper, I will only be concerned with moral responsibility as accountability. For the distinction between accountability and other notions or kinds of responsibility, see, for example, Watson (1996/2004), Shoemaker (2011), and Fischer and Tognazzini (2011).

2. For the notion of grounding, see, for example, Rosen (2010), Fine (2012), and the other chapters in Correia and Schnieder (2012). There is a debate about whether or not the grounding relation is transitive; see, for example, Schaffer (2012). The version of Strong Strawsonianism that I will defend does not depend on whether this is so.

3. Compare the claim that “facts [about whether people are morally responsible for what they do] can be interpreted as facts about whether it would be appropriate to hold people morally responsible” (Wallace 1994, p. 91). One could argue for such an identity thesis by claiming that it is hard to characterize a person’s being morally responsible for a bad action (in the sense of accountability) without at the same time characterizing her as being an appropriate target of blame for that action. This might be taken to suggest that these are the same properties. Examining the prospects of this line of reasoning is the aim of another paper.

4. Similarly, McKenna argues that Wallace’s account has difficulty dealing with the fact that our “norms and practices of holding responsible are sensitive to facts about agency that help settle whether an agent is equipped for being responsible. Often, our norms and practices bend to the nature of the agent, not vice versa” (McKenna 2012, p. 50). The following reply to the Naturally False Objection is also intended to work against this objection.

REFERENCES


