1. Consequentialism and moral influence

In the literature on free will and moral responsibility, there is an idea held by some that if we lack one or another form of libertarian agency (e.g., indeterministic event casual agency, or more traditionally, agent causation), all we have left are consequentialist justifications for praise and blame. This would be unfortunate, because the imagined consequentialist picture is by all accounts an impoverished one. It is typically taken to be a moral influence account, of the sort proposed by Moritz Schlick in his *Ethics*, which holds that moral responsibility is to be understood simply in terms of whether or not a given agent is susceptible praise and blame. The troubles with this sort of view are numerous and its critics are surely right that this comparatively unsophisticated approach to moral responsibility is ill suited to do the work of accounting for and justifying our complex web of responsibility practices. What all of this suggests is that moral responsibility would be in dire straits if we are not the sorts of agents described by libertarians about free will and moral responsibility. Call this the Dire Conclusion.

One of the aims of this paper is to undercut the Dire Conclusion by (1) presenting a picture of how a somewhat more plausible moral influence theory might be structured and (2) arguing that it need not be consequentialist. If I am right, it suggests that there is good reason to suppose that if we lack libertarian agency—as many suppose we do—this need not have the result that all we are left with are consequentialist accounts of moral responsibility, and maladapted ones at that.

A second aim of this paper is to comment on some complexities of the relationship of the theory of moral responsibility to normative ethics more generally. These remarks are intertwined with the discussion of moral influence theories throughout the paper.
Traditional moral influence accounts—such as the one proposed by Schlick—are often thought of as accounts of the content of responsibility norms, or sometimes as accounts of the nature of responsible agency. I agree with the consensus that a moral influence account is unworkable in these terms. Nevertheless, I think there is something right about moral influence having something to do with moral responsibility. In particular, I think we do better to think of moral influence as playing a role the *justification* of responsibility norms. The idea I will pursue here is roughly this: the justified responsibility norms are those norms of praise and blame that are most effective at, over time, getting creatures like us to attend to moral considerations and to govern our conduct in light of them. When construed in this way, the role of moral influence is at the level of systemic effects, rather than individual actions. First-order norms need not be understood in terms of influencing agents, nor need they to make reference to moral influence. What is required is that those norms—whatever they turn out to be—have the net effect of influencing the right sorts of agents in the right sorts of ways.

I will argue that this reconceptualization of the role of moral influence permits several innovations over the traditional and problem-ridden approaches. Perhaps more importantly they allow us to make sense of how we might justify our responsibility characteristic practices and attitudes (including praising and blaming) in more than consequentialist terms, and without appeal to libertarianism, if we must.\(^1\)

2. Two possible projects

---

\(^1\) Elsewhere, I have attempted to show the recast moral influence account I favor can answer the objections philosophers cite as reasons for rejecting traditional moral influence accounts. See [SUPPRESSED FOR REVIEW] . In the present paper, I will largely assume that traditional objections to moral influence accounts are irrelevant to the present theory, or that they can be addressed.
There are at least two possible tasks for a theory of moral responsibility, and we should be
careful to distinguish them. One task we could undertake is diagnostic. That is, we could offer an
account of the folk concept of moral responsibility. This task is common to conceptual analysis
methodologies in philosophy. A second task we could undertake is prescriptive. This account
would aim to tell us what we ought to believe about moral responsibility, even if it departed from
folk thinking about moral responsibility. Prescriptive accounts can have various degrees of
commitment to the folk concept or aspects of it, some enshrining certain linguistic or conceptual
intuitions as privileged, while others not. The important point is that, unlike conceptual analysis
accounts, prescriptive accounts are not limited to representing whatever concept or responsibility
we find ourselves in possession of.

What follows is prescriptive, directed at showing that a sizable subset of our
responsibility-characteristic practices can plausibly be justified in moral influence terms. It is not
concerned to characterize ordinary thinking about moral responsibility. Indeed, one of its upshots
is that even if we accept that the metaphysics of responsibility is somewhat different than folk
thinking makes it out to be, it need not lead to the destruction of our moral responsibility
practices.

---

2 While traditional forms of a priori conceptual analysis are not widely accepted, conceptual analysis in some form
or another has its staunch defenders. For examples, see Frank Jackson, From Metaphysics to Ethics: A Defense of
York: Oxford, 1992). The issue is somewhat complicated, however, by “paraphrasing” approaches to conceptual
analysis. For discussion, and for a related but somewhat finer-grained taxonomy of theoretical projects in the free
will debate, see Shaun Nichols, “Folk Intuitions on Free Will,” Journal of Cognition and Culture 6, no. 1 & 2
(Forthcoming in 2006).

3 On the diagnostic issue, I believe that the correct account of our folk concept of moral responsibility is
incompatibilist (i.e., incompatible with the thesis of determinism), and that — alas — our folk concept is unlikely to
be true. On the issue of prescription, I favor a moderately revisionist approach to things. I believe that much of our
moral practices and attitudes associated with moral responsibility can be justified on grounds independent of the
problematic (libertarian) folk concept, and that we should revise our concept of moral responsibility in light of what
is justified. Such a justification will minimally require revisions in our conception and theory of responsibility, and
perhaps our practices. The difference between commonsense thinking and what this account prescribes is what
makes the account I offer revisionist. What makes it moderate, is that it does not counsel full-scale elimination of
our responsibility-characteristic practices, attitudes, and beliefs. My account is therefore distinct from various
3. The justification of responsibility norms

A theory of responsibility norms is a theory that describes the norms of moral responsibility and their foundation. In contrast, a theory of responsible agency provides an account of when something is the kind of creature to which the norms of responsibility apply. It may be useful to discuss these issues in separation from one another, but it is important to recognize their interconnections. To know whether someone is responsible for some action or outcome, we must first determine whether that person is a responsible agent. Conversely, if we know that someone is a responsible agent we cannot settle issues of justified praise and blame until we know something about the norms of responsibility.

My concern here is primarily with a theory of the norms of responsibility. On my account, however, the norms of responsibility can only be properly understood by reference to an account of responsible agency. The account of responsible agency I favor is one that emphasizes an agent’s sensitivities to moral considerations. Roughly, an agent is a responsible agent (i.e., an appropriate subject of responsibility norms) in a particular context if the agent can detect and respond appropriately to moral considerations. The notion of capacity my account presupposes is compatible with our universe being deterministic, and thus the prescriptive account I offer can

---

skeptical versions of incompatibilism (i.e., hard determinism and hard incompatibilism) that do recommend that some or all of our responsibility-characteristic practices would need to be abandoned. I have argued for these views in a number of places. See, for example, [SUPPRESSED FOR REVIEW] I think of this paper as playing a role in justifying a revisionist prescription, showing how aspects of responsibility can remain well grounded in the absence of our libertarian self-conception.

4 The basic picture—responsible agency as crucially involving sensitivity to moral considerations—has been held by a number of authors, although the details vary by account. See, for example, Nomy Arpaly, Unprincipled Virtue (New York: Oxford, 2003), John Martin Fischer and Mark Ravizza, Responsibility and Control: A Theory of Moral Responsibility (New York: Cambridge University Press, 1998), Michael McKenna, “The Limits of Evil and the Role of Moral Address,” Journal of Ethics 2, no. 2 (1998), R. Jay Wallace, Responsibility and the Moral Sentiments (Cambridge, Mass.: Harvard University Press, 1994), Susan Wolf, Freedom within Reason (New York: Oxford University Press, 1990). I do not believe, as the authors of these accounts mostly seem to, that this general approach to accountability is non-revisionist.
be thought of as a kind of revisionist compatibilism. There is a great deal more that can be said about the nature of responsible agency, but for present purposes this should be sufficient.\textsuperscript{5}

Let us return to the idea of moral influence. The traditional model of moral influence got this much right: praising, blaming, punishing, and rewarding typically work on creatures like us by providing motivation for us to track moral considerations and to regulate our behavior in light of them. Traditional models erred, however, in holding that this was all responsibility consisted of. While it may be true that responsible agents are (typically) influenceable agents, it does not follow that moral responsibility consists solely in moral influence. A given agent might not be influenceable at all in a particular case (perhaps he is too angry or distracted or inebriated), but it is plausible to think that the agent might nonetheless count as a responsible, appropriately subject to the norms of responsibility. Moreover, a good many of our practices of responsibility holding are not performed with an eye towards downstream influence. I might feel gratitude for something you have done with no expectation of some future benefit. The content and rules of our practices are much more complicated than the simple picture proposed by moral influence theorists.

What traditional moral influence theorists got wrong, then, was either mistaking an accidental or occasional feature of our practice as the whole of the practice, or confusing the justification of our web of practices and attitudes with the content or imperatives of the responsibility norms. What I propose is to construe moral influence as characterizing the structure of justification of a complex set of practices, and not a description of the particulars of every instance of justified praise and blame. We can put the idea this way: \textit{what justifies the bulk of our responsibility-characteristic practices and attitudes is that this system tends to, over time, and given constraints of stability and the limits of our psychology, get creatures like us to attend

\textsuperscript{5} I have offered my own account of responsible agency in [Suppressed for review]
to moral considerations. Although we need not assume our current practices are maximally efficient at doing this, it is reasonable to suppose that at least in broad strokes, our responsibility-characteristic practices, attitudes, and beliefs do constitute a system that tends to, over time, get many of us to attend to what moral considerations there are.

There are at least four aspects of my reformulation of the moral influence idea that I will spend a few moments unpacking. These ideas include the view’s approach to indirectness, psychological complexity, inflexible reactive attitudes, and revisionism.

First, indirectness: the proposed reformulation treats moral influence as a second-order property of a system of responsibility norms. It is the proper aim of a system of responsibility norms, but the individual norms themselves need not make any mention of influence. As with many social practices, the aim and the practice can diverge in interesting ways. Whatever the aim of dating may be, it is typically a bad strategy to pursue that aim directly in every dating-related interaction. An effective set of practices will often contain large webs of activities or norms that do not, considered in themselves, directly contribute to the attainment of the goal. They may nonetheless indirectly contribute to the goal, and sometimes, essentially so.⁶ What matters is that the system is efficacious as a whole, over time, with creatures like us.

None of this is to deny that there may also be justifications for aspects of our responsibility-characteristic practices, attitudes, and beliefs that do not turn on their role in a system of responsibility directed at fostering our sensitivity and responsiveness to moral considerations. For that matter, expression of the responsibility-characteristic attitudes and practices need not be exclusively tied to their characteristic role in moral responsibility. Blaming

---

⁶ This point has been made in a number of ways in the context of consequentialist moral theories. See, for example, Robert Merrihew Adams, “Motive Utilitarianism,” Journal of Philosophy LXXIII (1976), Peter Railton, “Alienation, Consequentialism, and the Demands of Morality,” Philosophy and Public Affairs 13 (1984). This idea is also common in the literature on decision theory.
can express solidarity in some shared value, commitment to some social norm, or admonishment or encouragement for others to behave in a particular way or to become a particular kind of person. My reformulation is therefore non-exclusive about the justification of our responsibility-characteristic practices, attitudes, and beliefs, and compatible with blaming that has motivations beyond simple responsibility ascriptions. For that matter, the psychological mechanisms involved in holding ourselves and others responsible may play other roles in our psychological economy, above and beyond the role they play in the context of moral responsibility. This may mean that our psychology is not optimized for smoothly translating responsibility norms into moral motivation. It may also mean that our practices will have a certain degree of looseness built in to them, a degree of flexibility that reflects the ways in which our psychology of recognition and motivation are constrained by their structure and varied aims.

These considerations point to a second important aspect of the account I am offering: we need not assume a crude psychology of responsibility ascriptions, where all of us are concerned to praise and blame one another with an eye towards influencing future behavior. We can and should suppose that we have diverse concerns in our ascriptions of responsibility. Sometimes we may praise and blame with an eye towards the future and other times not. Indeed, some agents may never be thinking about responsibility in future-directed terms. On this account, then, there is room for both backward and forward-looking considerations to play a role in our ascriptions of moral responsibility.

A third important aspect of the present account is that it compatible with the potential discovery that some of the psychological mechanisms that the system of responsibility relies upon—including what P.F. Strawson called “the reactive attitudes”—will prove to be inflexible and perhaps even impediments to specific or general practices that over time tend to get
responsible agents to track moral considerations. Perhaps, on balance, resentment hinders sensitivity to moral considerations more than it fosters attention and response to them. If so, and if resentment is a non-negotiable aspect of our cognitive and affective architecture, then any plausible theory—including a reformulated theory of moral influence—will acknowledge these features as constraints around which the theory must be shaped.

Whether an agent is morally responsible for his or her actions is not a function of that particular agent’s susceptibility to influence, but rather a function of whether the agent is a responsible agent, and if so, what the justified norms of moral responsibility say about responsible agents in those contexts. What makes responsibility norms justified is that they tend to, over time, influence moral consideration-responsive creatures like us to better attend to what moral considerations there are. These norms will tend towards generality and not to act or token-specific norms of responsibility. This tendency will be driven by the realities of our psychologies, including the length of time it takes to develop and refine moral attitudes, the flexibility of our attitudes, the cognitive burden involved in assessing responsibility, and the overarching need to have a stable and efficacious system of influence. To put the point differently, what facts there are about the norms of responsibility will be constrained by the kinds of creatures we are and the aim of getting us to be better dealers and consumers in the trade of moral considerations.

A fourth and final aspect of this account merits note; the account is compatible with revisionism about the folk concept of moral responsibility. As a prescriptive account, it is not intended to capture the full contents of ordinary thinking. Indeed, I am inclined to think there are various difficulties with the metaphysics of commonsense conceptions of responsibility, and the present proposal can fairly be construed as a blueprint for reconstructing our thinking about
responsibility in light of those difficulties. Although my concern here is to emphasize how this reformulation of the idea of moral influence might justify a large subset of our responsibility-characteristic practices, it is an open question whether there are better ways to achieve the aims of moral influence that would involve something more than the comparatively conservative restructuring of our concepts that I am proposing.

Revisionist proposals can raise the worry that the theory is a shifting target, where any objection is met with the “but I’m a revisionist!” reply. These concerns are legitimate, but they can also be addressed. First, a revisionist account is only well suited to ignore counterintuitive results that rest upon some aspect of commonsense that is getting revised. These results can be ignored precisely because the aim of the theory is to change that aspect of common sense. However, when the counterintuitiveness of the theory or its results does not turn on the issue or concept being revised, this should prove to be as worrisome as counterintuitive results in any domain. Second, even accounts that make prolific use of revisionist maneuvers will still have to be weighed against alternative accounts. Other things being equal, it is reasonable to suppose that a more conservative revisionist account will be preferable to more radical versions of revisionism, given the high degree of difficulty in instituting any widespread conceptual change—especially a philosophical one. So, even if there is reason to adopt revisionism in some domain, there will always be pressure to favor the least radical revisionist account that does the work that cannot be done by a non-revisionist account. Although more might be said about this, we might summarize things in the following way: Departures from commonsense are not troubling if either of two conditions hold for the considered case: (1) commonsense relies on a metaphysically implausible notion of agency or (2) we can offer a revisionist account of some
considered issue that is on balance better than the account presupposed by commonsense. For these reasons, revisionism does not, by itself, guarantee that the theory is immune to refutation.

4. Modularity

It is a shortcoming of contemporary accounts of moral responsibility and normative ethics respectively that they each tend to say precious little about their relationship to one another. If one simply followed the literature on moral responsibility, one could be forgiven for thinking that moral responsibility had nothing whatsoever to do with morality. If one followed the ethics literature, one could be similarly forgiven for thinking that moral responsibility was no part of ethics, or at any rate, that there are no interesting philosophical or ethical issues raised by the notion of moral responsibility.

One aspect of my recasting of the moral influence approach to moral responsibility is that the account turns out to be modular, compatible with a range of normative ethics. That is, my account of moral responsibility describes the internal logic and justificatory structure of our responsibility-characteristic attitudes, practices, and beliefs, but in a way that is compatible with at least many approaches to normative ethics. Indeed, the account importantly depends on some larger, background account of moral norms to fill in some of the details of a theory of responsibility. Depending on the account of normative ethics, aspects of the responsibility module will take on a different character.

Call the correct theory of normative ethics, whatever that theory turns out to be, the True Theory. The True Theory completes some of the details of a modular theory of responsibility by, for example, giving us a better sense of what constitutes a moral consideration. The True Theory

---

7 There are a handful of exceptions, the most well known of which is probably T.M. Scanlon’s contractualist account of these things. See T. M. Scanlon, What We Owe to Each Other (Cambridge, Massachusetts, and London, England: The Belknap Press of Harvard University Press, 1998).
may also interpose considerations, external to the norms of responsibility, which mitigate, counterbalance, or trump the edicts of the responsibility norms. Embedding a modular account of responsibility into the True Theory both adds detail to the account of moral responsibility and fights the myopia that can come from focusing on moral responsibility in isolation from normative ethics.

(This is not to prejudice the issue of direction of theory influence. We might find that considerations from normative ethics affect what theories of moral responsibility we find plausible. We might also find that the reverse is true: considerations native to the theory of responsibility might favor or disfavor particular ethical theories.\(^8\) The same might also be said of various metaethical considerations. For example, some philosophers have thought it important to account for the apparent fact of the matter whether someone is responsible. Considerations like these might provide some impetus in favor of cognitivist accounts of moral language.\(^9\)

If we accept the modularity thesis, it should become clear that a sophisticated moral influence theory need not be tied to consequentialism. It could be, but this is not an essential or intrinsic part of a moral influence account of moral responsibility. Although the justification is teleological (aimed at fostering a kind of agency suitably responsive to moral considerations), and as such might be viewed as a kind of consequential reasoning, it does not follow that one’s normative ethics must be consequentialist. Indeed, it bears recalling that the norms themselves are not particularly consequentialist in their content. What makes you responsible in a particular instance is not whether or not holding you responsible gets the best results with respect to fostering responsiveness to moral considerations. What makes you responsible is determined by whether one is the right sort of agent (i.e., sensitive to and capable of guiding your conduct in

\(^8\) Thanks to Andrei Buckareff for bringing this point to my attention.

\(^9\) [Suppressed for review]
light of moral considerations) in the relevant context and what the norms of responsibility say about agents in those circumstances. As we saw above, the content of responsibility norms might have little direct concern for consequences, maximization, or other marks of a straightforwardly consequentialist ethic.

Kantians and proponents of other approaches to normative ethics need not forgo a moral influence account of the justification of responsibility simply because core aspects of their normative ethics do not rely upon consequentialist considerations. That I think right action consists in obeying the categorical imperative does not mean I cannot think there are ever issues where consequences are of concern. Indeed, I might think that given the dignity and intrinsic value of rational, self-controlled creatures it would be an entirely appropriate thing for us to have a system of norms and associated practices whose concern was precisely the fostering of that form of agency. To fail to endorse and participate in this practice would show a failure to value highly enough the kind of agency at which this practice is directed at developing, a kind of agency I value in myself and ought to value in others.

Indeed, embedded in a non-consequentialist account of normative ethics, a sophisticated moral influence theory gains the benefit of various non-consequentialist conceptual resources (rights, moral character, duties, obligations, practical wisdom), whose independent standing can fund replies to a variety of worries that have been raised against traditional, non-modular moral influence accounts. Traditional moral influence accounts face the worry that there seem to be a range of cases where we could hold someone responsible and derive significant downstream benefits from doing so, even when the agent was not (intuitively speaking) morally responsible. Scapegoating—for example, executing an innocent man to appease an angry mob—or certain practices of strict liability have this structure. Conversely, there are cases where an agent seems
morally responsible but where there are not downstream benefits to be had. We have already seen how the proposed account handles the latter case; whether or not someone is responsible is settled internal to the norms of responsibility (including norms of responsible agency), and those norms need not be concerned with downstream effects. The former case, where there are benefits to holding someone responsible who does not seem responsible, is handled both similarly and differently. With respect to the issue of whether the agent is responsible, this is settled by the norms of responsibility and the conditions of responsible agency. With respect to whether it is permissible or desirable to falsely treat someone as responsible turns on the theory of normative ethics in which the account of responsibility is embedded. If the True Theory rules out scapegoating (for example, because it uses someone merely as a means) then it will already be ruled out for our practices of responsibility holding. If, however, the True Theory permits it, then if it is a problem, it is not so much a problem for the theory of moral responsibility so much as it is a problem for the True Theory.¹⁰

Similar things might also be said of the general worry that the justification of a system of norms does not tell us what to do in a particular case. In the case of a theory of responsibility, these considerations are somewhat dampened by the need to have a workable set of generally applicable principles that can guide our conduct in a way that respects and relies upon widespread features of human psychology. These are the pressures to generality and stability I mentioned in section 3. In the context of a modular theory, this concern is further dampened by the recognition that we have a wide range of ethical concerns beyond moral responsibility. These ethical concepts can and should inform our decisions about particular cases, and we must allow

¹⁰ I am, however, inclined to think that there are constraints native to the logic of responsibility norms that will constrain scapegoating and some versions of strict liability, since both of these practices seem to largely bypass the mechanisms of tracking and responding appropriately to moral considerations. I will not try to make that argument here, however.
for the possibility that considerations grounded in other aspects of ethics might guide, mitigate, counterbalance, or trump what is prescribed internal to the norms of responsibility.  

One might object that an upshot of the kind of modularity I have described is that a fleshed out account of the responsibility module — considered by itself — would remain silent on whether or not you are to praise or blame someone. Although it may well provide an account of whether or not someone is responsible and it might well provide an account of what justifies responsibility ascriptions in general, what it does not do is provide an account of what one should do in a particular concrete case, as what it provides are roughly rule-like admonishments whose application may fail in a given instance.

I am inclined to agree that an account along the lines of what I have offered is only partially action guiding. The extent to which it is action guiding is this: the account does say that ordinarily one is permitted to blame when someone is blameworthy and praise when praiseworthy. More specifically, there will be pressure for these sorts of general rules of thumb to hold in unexceptional circumstances: inasmuch as praising and blaming tend to keep us sensitive to moral considerations, and inasmuch as these practices are parasitic on a messy, organic psychology, the costs (including the cognitive load) of determining whether one should violate a generally good rule of thumb will generally weigh in favor of simply accepting the rule of thumb and acting on it. One might protest that this still leaves some gap between what the account of responsibility norms recommends and what an individual should do. If I am right, however, this gap cannot be closed by a theory of responsibility by itself. What is needed to fill that gap is some more general account of moral or practical considerations that either trump or

---

11 There is, of course, the further issue about the extent to which generally justified norms are genuinely considerations for me. Again, these issues are complex and go far beyond the theory of moral responsibility, so I will not attempt to say anything about them here.

12 Thanks to Caspar Hare for having pointed this out in conversation.
finish elaborating the details of a theory of responsibility norms. Without leaving this sort of space we would be left with the strange result that the one place a theory of ethics must fall silent is on issues of moral responsibility.

All of this shows that the moral influence theorist need not accept the charge that moral influence accounts trample an important distinction between when someone is responsible and whether it is appropriate to hold him or her responsible. Indeed, this distinction is part and parcel of any recognition that while a system of moral responsibility has something of an internal logic to its norms, it is nonetheless part of a broader system of normative ethics. So, although a moral influence theory may be mated to a consequentialist account of moral responsibility it does not need to be. Indeed, coupling the account I have offered with a non-consequentialist account of normative ethics might be the right mix of revision and conservation, at least for purposes of finding a point of wide reflective equilibrium.

5. Conclusion
In this paper I set out to explore some of the connections between a plausible theory of responsibility and normative ethics. I also aimed to challenge the claim that if we lack libertarian agency, we are left with an impoverished and largely unworkable consequentialist account of moral responsibility. My approach has been to present a picture of how a workable moral influence theory might be structured and to consider the extent to which this sort of picture must be wedded to consequentialism. I have argued that a sophisticated moral influence picture should be thought of as an account of the justification of a network of our responsibility-characteristic practices and attitudes, and not as an account of the justification or operation of individual instances of responsibility holding. If I am right, this account functions without appealing to a
libertarian picture of agency. Whatever benefits libertarianism might bring, it is not required to justify our praising and blaming of each other in the ways characteristic of moral responsibility. Moreover, the sort of justification I have offered does not collapse into a purely consequentialist justification for praise and blame. The account is compatible with aspects of our responsibility practices being justified in other terms, and the terms in which I have justified the bulk of our practices is compatible with non-consequentialist moral theories. What my account does require is that there be something valuable or worthwhile about a kind of agency that is sensitive to and appropriately responsive to moral considerations. It is surely plausible to think that there is something valuable about this form of agency, and although I have not tried to defend it here, we plausibly have this sort of agency with some frequency. If so, then regardless of how debates turn out for the compatibility of, on the one hand, our folk concepts of responsibility and free will, and on the other, determinism or naturalism, we can be confident that the justification and normative integrity of responsibility practices, per se, is not at stake.
Works cited


