

Moral Necessity and the Implementation of Modernity

blind review

1 Introduction

There are two competing views on moral responsibility. One view gives priority to *conscience*, an essentially (or existentially) human feature calling individuals to define themselves and perform actions in keeping with a particular way in which human beings understand themselves apart from or in addition to what it means to be human in scientific or rational ways. Even though it is tempting to attribute this view more broadly to continental philosophy rather than analytic philosophy, I have selected a continental and an analytic philosopher to represent it: Martin Heidegger and P.F. Strawson.

Another view embeds moral responsibility in the way in which we more generally view decision-making and motivating behaviour. Moral agents have distinctly moral ends (such as happiness for sentient beings or human flourishing) to which the performance of morally motivated acts is the means. Metaphysical questions as to what is or what is not moral often retreat into the background for these accounts of moral responsibility, whereas questions of rationality are foregrounded. A moral agent must to some extent be epistemologically empowered in the sense that her internal deliberations, intentions, choices, actions, and the resulting outcomes are suitably aligned.

Even though it is tempting to attribute this view more broadly to analytic philosophy (see, for example, the prominence of utilitarian ethics in analytic philosophy), my view is that there is a strong contingent in continental

philosophy taking up this viewpoint as well. Its approach, however, is often critical of the possibility that human nature or the human condition is meaningfully consistent with the requirements of morality. Some continental thinkers follow the mainstream of analytic philosophy that “reactive attitudes remain within the bounds of reason” (Russell, 1992, 302), but add the element that human beings are not in the least equipped to meet standard rationality requirements.

Some continental thinkers have thought through the implications of what it means that on the one hand humans consider themselves morally responsible and on the other hand scientific and technological progress has put them in an unprecedented situation of epistemological malaise with respect to understanding themselves. I have selected Franz Kafka and Michel Foucault as representatives of this view; neither of them analytic philosophers (one of them not a philosopher at all), but it is my intention to show how their position can be rendered intelligible in the terms of analytic moral philosophy and what the consequences are for moral philosophy in general.

2 Implementation of Modernity

In his article “Freedom and Resentment,” Strawson embarks on a mission in moral philosophy that runs parallel to Jürgen Habermas’s idea of detranscendentalizing in the philosophy of rationality. Immanuel Kant had given both the philosophy of rationality and the philosophy of morality strongly transcendental credentials, and both Strawson and Habermas are committed to saving the enlightenment project from what Strawson calls ‘over-intellectualizing.’

Strawson’s argument proceeds as follows. There are optimists and pessimists about the role of determinism for moral responsibility. The optimists (compatibilists) do not consider pronouncements of whether the world is or is not deterministic a threat to moral responsibility. If anything, determinism supports the efficacy of making moral distinctions because it helps to make concepts such as deliberation, intention, and planning intelligible.

The problem, as the pessimists (libertarians, skeptics) point out, is that moral distinctions would then fail to be merit-based, as effective as they may be. Morality would devolve into “intellectual understanding, management, treatment, and control” (Strawson, 2008, 18). For the pessimists, this devolution constitutes a failure to give a coherent account of what is meant by morality. Kafka and Foucault, who are Strawsonian pessimists of sorts, consider this devolution to be representative for the implementation of modernity, but more about this in a moment. Strawson’s solution to the pessimist’s challenge is to detranscendentalize morality. The pessimists have to give up on their stilted metaphysics; the optimists have to give up on the moral picture “painted in a style appropriate to a situation envisaged as wholly dominated by objectivity of attitude. The only operative notions invoked in this picture are such as those of policy, treatment, control” (Strawson, 2008, 22).

It is central to Strawson’s argument that moral judgments are not primarily objective but rather participant attitudes. Humans have a natural commitment to ordinary interpersonal attitudes. “This commitment is part of the general framework of human life, not something that can come up for review” (Strawson, 2008, 14). According to Strawson, we cannot take seriously the thought that a high-level theoretical question such as whether or not determinism is true would lead to the decay of the personal reactive attitudes. Jonathan Bennett summarizes,

if we try to imagine our lives without reactive feelings we find ourselves . . . confronted by bleak desolation. We cannot be obliged to give up something whose loss would gravely worsen the human condition, and so reactive feelings cannot be made impermissible by any facts. (Bennett, 1980, 29)

In keeping with a cultural trend in the philosophy of the mid-20th century, Strawson’s account of moral responsibility is anti-metaphysical (consider, for example, the skepticism towards metaphysical questions in and around the Vienna Circle). Moral responsibility can be preserved from the threat

of an invasive scientific anthropology: at the expense of metaphysics and in favour of, as Habermas calls it, a situated view of humanity. For Habermas, transcendental rationality gives way to a rationality implemented in communicative action that aspires towards an ideal speech situation. For Strawson, the cold-hearted scientific approach of the optimist and the metaphysical exuberance of the pessimist give way to a realistic and embodied assessment of human psychology.

What violates this assessment is the idea that moral agents can fully emancipate themselves from their participant attitudes and hold objective attitudes instead. Strawson goes into some detail what the dystopian consequences of a moral account would be that trades solely in objective attitudes (examples may be the Kantian moral agent whose moral sense is derivative of transcendental rationality; or the utilitarian who performs a calculus to identify the optimal means to achieving her moral aims).

A sustained objectivity of inter-personal attitude, and the human isolation which that would entail, does not seem to be something of which human beings would be capable, even if some general truth were a theoretical ground for it. (Strawson, 2008, 12)

Strawson goes on to describe the ‘abnormal egocentricity’ required to maintain that a metaphysical theory (such as determinism) can lead to the collapse of moral structure in human behaviour and motivation. Both human communities and individual personhood are based on participant attitudes, therefore circumstances in which everyone is morally incapacitated as a matter of metaphysical theory would lead to a breakdown both on an individual and a social level.

My claim is that there is a sense in which Kafka and Foucault take up this challenge and show in which ways it is not only possible that Strawson’s dystopia is implemented but that modernity has already largely progressed to a state where both individuals and collectives no longer meaningfully identify as moral agents.

Strawson has confident pronouncements about what is and what is not essential to human nature. One might claim that there will never be an end to poverty (Jesus appears to do so in Matthew 26:11), tribalism, jealousy, sexual harassment in the workplace, or violence. Claims of this sort appeal to human nature. For Kafka and Foucault, the problem is primarily not that some human traits are genetically fixed or metaphysically ordained, but that if anything the tendency is to underestimate the ability of a human being to alienate herself from herself. Both Kafka and Foucault are to some degree epigones of Sigmund Freud in this instance, who was not chiefly interested in pathological *expression* of jealousy and violence, but in their pathological *suppression*. For Kafka and Foucault, the alienation of modern humans from their attachment to moral responsibility becomes an issue and a (grim) possibility.

In his book *Being and Time*, Martin Heidegger argues that the distance created by the existence of a human being and the fact that this existence is an issue of concern for that human being becomes the foundation for existence as 'Dasein.' Dasein has a different ontological structure than the existence of galaxies, chairs, the Higgs boson, black swans, or unicorns. Moral responsibility is thus immediately divorced from mere facts about external reality. In §58, Heidegger characterizes moral responsibility in terms of conscience, a call by the self to the self towards the possibilities of Dasein.

What Strawson and Heidegger have in common is the necessity of de-intellectualized morality for human life. For Strawson this is a simple fact of human nature (one that may, for example, be open to scientific inquiry). For Heidegger, it is a more fundamental condition for existence as such. Guilt (one finds echoes of this view in Kafka's work) is primarily an intransitive ontological condition rather than a transitive consequence of the distinction between good and evil. Transitivity is the grammatical feature of requiring an object, in this case an object of guilt, the violation or transgression of a rule.

In the following, I will discuss pushback against this insistence on necessity for moral responsibility. In his account of justice in *A Treatise of Human*

Nature (3.2.1), David Hume provides one of the first instances of what Friedrich Nietzsche would later call a ‘genealogy.’ In this account, Hume relativizes the metaphysical credentials of justice and characterizes it instead as organically grown in the soil of social interaction among humans. It is critically important to Hume, however, that justice arises from the social needs of humans as a matter of necessity.

Foucault describes in his work, whether it is about the hermeneutics of the subject, the penal system, schools, hospitals, prisons, or the history of sexuality, that there is nothing necessary about the particular forms that moral explanations take. They are rooted in contingencies and are in the final analysis a downstream consequence of the implementation of modernity (for Foucault, especially the economic, institutional, and bio-political changes that modernity brings about).

We now have a set of strong contrasts. The optimists versus the pessimists of Strawson’s argument; Strawson’s psychological assumptions about human nature versus Heidegger’s existential assumptions about the human condition; and finally, accounts of moral responsibility that consider its most fundamental expressions to be necessary to human life versus accounts which attribute to moral responsibility ineradicable contingency.

In an obscure French essay about Kafka, Claude-Edmonde Magny makes a compelling observation:

One can find in Kafka’s work a theory of responsibility, views on causality, finally a comprehensive interpretation of human destiny, all three sufficiently coherent and independent enough of their novelistic form to bear being transposed into purely intellectual terms.¹

¹The translation of the quote is from Blanchot, 1995, 2, the original quote is “on peut trouver dans l’œuvre de Kafka une théorie de la responsabilité, des vues sur la causalités, enfin une interprétation d’ensemble de la destinée humaine, suffisamment cohérentes toutes trois et assez indépendantes de leur forme romanesque pour supporter d’être transposées en termes purement intellectuels” (see Magny, 1945).

This paper, which is committed to the view that there is philosophical insight in both the analytic and the continental tradition, has more specifically two aims: to shed a critical light (i) on Strawson's account of moral responsibility, especially its claims about necessities of human nature; and (ii) on Kafka's ability to address in literature philosophical claims that can be rendered analytically intelligible. Kafka is a pessimist vis-à-vis the type of moral agency envisioned by Strawson. I will try to detail this pessimism in the next section.

3 Kafka's Account of Moral Responsibility

There is a view that Kafka's work is elusive with respect to interpretation. Various schools of interpretive approaches to Kafka have formed: Marxist, existentialist, psychoanalytic, and others. Most commentators, however, claim some place on the sidelines of the controversy between these schools and underline how open Kafka's texts are to all sorts of interpretation. W.H. Auden, for example, states in the first sentence of his essay on Kafka "The I Without a Self" that Kafka is the master of a literary genre "about which a critic can say very little worth saying" (Auden, 1962, 159). For this view, it is the fluidity of Kafka's texts by which they effortlessly move from being viewed from one perspective to being viewed from an entirely different perspective that constitutes Kafka's genius.

My claim, by contrast, is that there are themes that run through Kafka's literary output which can be captured by interpretation, translated into the terms of analytic philosophy (no doubt there will be losses in this translation, losses nonetheless worth the rewards), and used to manifest a definitive position that Kafka takes with reference to topics such as moral responsibility, epistemological access to causation, the status of metaphysical beliefs, and the meaning of happiness.

There is an aporia in human life: we hold people morally accountable for their actions and choices, yet we are also aware that human decision-making

proceeds in a rich and multi-layered field of causal influences. Moral blame attenuates as a superficial investigation yields to a deeper one. In the *Nicomachean Ethics*, Aristotle carefully distinguishes between different types of ignorance, especially ignorance about one's true interests on the one hand and ignorance about particular circumstances on the other hand. The former ignorance characterizes vice, whereas the latter may annul it. A writer like Kafka revels in the description of moral narratives that lead these kinds of distinctions ad absurdum.

Kafka is not only a Strawsonian pessimist in the sense that the metaphysics of modernity results in the devolution of moral coherence; he is also an epistemological pessimist. He is not a skeptic in the sense that nobody knows anything. Rather, the moral agents in Kafka's work know and are deceived about a wide variety of things of great importance to them, and as they investigate them (more often than not in good faith), it becomes apparent that their experience of epistemological failure in the past forestalls any epistemological confidence in the future.

While moral agents may know things, they can never be sure of things or trust their sensitivities. Knowing things does not translate into knowing that you know them. This view is incompatible with knowledge internalism, where $K(X)$, knowing X , implies $K(K(X))$, knowing that one knows X . The moral agent hardly ever has conscious access to the causal mechanisms that inform morally responsible decision-making. What was yet grotesque to Aristotle, that an agent should be ignorant of swaths of relevant circumstances, particularly his or her own identity ("now no one, unless mad, could be ignorant of all these circumstances together; nor yet, obviously, of the agent—for a man must know who he is himself," *Nicomachean Ethics*, 1111a.1), turns into the cultural malady of an epoch.

Everything of which we become conscious is arranged, simplified, schematized, interpreted through and through—the actual process of inner 'perception,' the causal connection between thoughts, feelings, desires, between subject and object, are absolutely hidden

from us (Nietzsche in *The Will to Power*, section 477).

I will use Nietzsche's and Foucault's idea of genealogy to show that it is constitutive of considering oneself a moral agent that the origins of moral agency remain masked for the moral agent. An unmasking of a genealogical concept such as moral agency or responsibility results in its collapse. In the context of Hume's account of justice, genealogy denotes the kind of explanation pointing to the origins of a social practice of which it is essential that they themselves are not used as reasons to follow the practice. The core of the practice is somehow constituted by a certain forgetfulness toward its history. The forgetfulness is at the root of lending the practice intrinsic rather than instrumental value. The intrinsic value becomes detached from the original usefulness of the practice. It is a value which experiences a threat to its reflective stability, and possibly a breakdown, when its historical origins are uncovered.

Hume considers the emergence of justice from this genealogical process to be a necessary and positive development—Nietzsche and Foucault, by contrast, develop a critical theory of genealogy and strip it of all elements of necessity. They give genealogical accounts of morality and truth (Nietzsche), knowledge and sex/gender (Foucault), meaning to undermine the intrinsic power of these concepts to motivate behaviour.

Deception is then not derivative of epistemological clarity but the other way around: there is at first a deception needed to get moral responsibility off the ground, namely that the diffuse and contingent origins of moral agency are concealed from the moral agent; only then can the moral agent come to some clarity of what her alternatives are and how these may play out in the future. The agent then makes a moral choice (such as Kafka's K. leaving Frieda to spend time with the family of Barnabas in *The Castle*), which is then vitiated by future contingencies and future diffusions.

At the end, there is the existential guilt that Martin Buber has identified in Kafka's texts (see Buber, 1960): a guilt which is purely formal and lacks all content, for while humans live under the burden of moral culpability, they

cannot legitimately be made responsible for anything in particular. Kafka expresses this repeatedly in the *Letter to His Father* when he identifies a “boundless sense of guilt” (Kafka, 1966, 68) while insisting that both he and his father are “entirely blameless” (loc. cit., 4) and declaring the “guiltlessness of us both” (loc. cit., 100).

Adam Smith, Charles Darwin, and Karl Marx consider an invisible hand at work in history. Despite its diffusion (which, more positively, can be cast as its essentially democratic core) and its lack of conscious intention on the part of an author or creator it contributes towards progress: in capitalism, the invisible hand creates welfare; in evolutionary theory, it creates and multiplies life; and in Marxism, it will not fail to bring about revolution and the victory of the proletariat.

Nietzsche, Freud, and Foucault provide a more pessimistic picture of the invisible hand. Foucault, for example, describes an intricate network of micro-dominations which concatenate to produce technologies of power that victimize without perpetrator. Sigmund Freud’s invisible hand operates during an individual’s ontogenesis (rather than Darwin’s phylogenesis) in early childhood, where the origins and explanations for these operations are wiped out by amnesia, which makes the history of human agents inaccessible to them. The invisibility of this hand is everything but benign, leading to widespread psychological pathology.

Nietzsche produces an account of concepts—concepts which to most of us sound metaphysically and eternally established—that locates their origins in the contingencies of human history: concepts such as morality, love, justice, knowledge, and truth. Bernard Williams describes Nietzsche’s idea of genealogy as follows:

A genealogy is a narrative that tries to explain a cultural phenomenon by describing a way in which it came about ... our ethical ideas are a complex deposit of many different traditions and social forces, and they have themselves been shaped by self-conscious representations of that history. However, the impact of these historical processes is to some extent concealed

by the ways in which their product thinks of itself. (Williams, 2004, 28)

Foucault takes Nietzsche's ideas further and explains:

However, if the genealogist refuses to extend his faith in metaphysics, if he listens to history, he finds that there is 'something altogether different' behind things: not a timeless and essential secret, but the secrets that they have no essence or that their essence was fabricated in a piecemeal fashion from alien forms. (Foucault, 1977, 142)

Kafka anticipates many of Foucault's conclusions in his literature. The enlightenment project of characterizing human life as guided by reason and by the transparency of self and nature to the third-person gaze of a human mind has failed. In his texts, Kafka carries on a tradition that is discernible already in Fyodor Dostoyevsky's *Notes from the Underground*:

Before your eyes the object vanishes, the reasons evaporate, the culprit is not to be found, the offence becomes not an offence but a *fatum*, something like a toothache, for which no one is to blame ... you haven't found the primary cause.² (Dostoyevsky, 2004, 18)

Transitivity gives way to intransitivity. Objects vanish, as in Martin Heidegger's fundamental ontology hermeneutics becomes operative on 'being' without a text to serve as an object for the hermeneutics. Kafka expresses this idea in literature by recording no particular work to do for the officials in *The Castle* and no object for Joseph K.'s guilt in *The Trial*. There is also no sovereign chooser of moral value as there is in existentialism. Perpetrator-subjects vanish in favour of bureaucratic concealment, as in Foucault's modern penal

²Compare also Émile Zola's character La Maheude in *Germinal* exclaiming, "Mais on réfléchit, n'est-ce pas? On s'aperçoit qu'au bout du compte ce n'est la faute de personne ... non, non, ce n'est pas ta faute, c'est la faute de tout le monde," page 497; and Martin Heidegger's dictum in *Sein und Zeit*, "In der Alltäglichkeit des Daseins wird das meiste durch das, von dem wir sagen müssen, keiner war es," section 27. Both of these quotes unfold their meaning specifically in the context of modernity.

system. Again, Kafka reflects this in *The Castle* and *The Trial*. Finally, moral responsibility collapses because there is for it no epistemological access to what the relevant causal relationships are: “no one is to blame . . . you haven’t found the primary cause.”

In Kafka’s *Metamorphosis*, Gregor “suffers in person from [the] evil consequences [of his employment], which he can no longer trace back to the original causes” (Kafka, 1995, 83). In *The Castle*, K., “in order to obtain pardon, first had to establish guilt, and that’s precisely what they denied him at the offices” (Kafka, 1998, 214). In the short story “On the Tram,” Kafka gives expression to the epistemological crisis that precipitates a moral crisis:

I stand on the end platform of the tram and am completely unsure of my footing in this world, in this town, in my family. Not even casually could I indicate any claims that I might rightly advance in any direction. I have not even any defence to offer for standing on this platform, holding on to this strap, letting myself be carried along by this tram, nor for the people who give way to the tram or walk quietly along or stand gazing into shop windows. Nobody asks me to put up a defence, indeed, but that is irrelevant.

In another short story called “Resolutions,” Kafka concludes:

So perhaps the best resource is to meet everything passively, to make yourself an inert mass, and, if you feel that you are being carried away, not to let yourself be lured into taking a single unnecessary step, to stare at others with the eyes of an animal, to feel no compunction, in short, with your own hand to throttle down whatever ghostly life remains in you, that is, to enlarge the final peace of the graveyard and let nothing survive save that.

4 Conclusion

Moral decisions that human beings make in relationship with each other are comparable to the decisions a driver has to make when steering a fast

vehicle without brakes in the fog. Kafka is not a nihilist—decisions matter. The driver in the fog makes consequential decisions, but the driver does not have enough information to make these decisions so that they become appropriate objects of moral evaluation.

Kafka had particular philosophical ideas which he used to give shape to the literature that he created. He was skeptical about accounts of moral responsibility produced by the enlightenment, such as Kant’s critique of practical reason or J.S. Mill’s consequentialism, because these moral theories make epistemological assumptions that are not borne out in human life. They fail by the standards of descriptive moral theory, which for both Kafka and Foucault is importantly primary to prescriptive moral theory. Descriptive moral theory is interested in the ways in which our behaviour is morally motivated and how these motivations interact with institutions, interests of power, the economy, contingencies of culture, and the management of bodies.

For Strawson, one way in which philosophers (both optimists and pessimists) have erred is that they have subordinated prescriptive moral theory to a particular type of description, the metaphysical theory of determinism. On the one hand, if Strawson is correct and it is in principle not possible to undermine fundamental moral commitments by high-level descriptions, then Kafka’s and Foucault’s project has failed. On the other hand, if Kafka and Foucault succeed in providing a persuasive account (the one a literary account, the other a philosophical account) describing just the kind of human isolation and moral alienation that Strawson considers inconceivable, then a weakness in Strawson’s argument is revealed.

References

Auden, W.H. *The Dyer’s Hand and Other Essays*. New York, NY: Vintage Books, 1962.

Bennett, Jonathan. “Accountability.” In *Philosophical Subjects*, edited by

Zak van Straaten, Oxford UK: Clarendon, 1980.

Blanchot, Maurice. *The Work of Fire*. Stanford, CA: Stanford University, 1995.

Buber, Martin. “Guilt and Guilt Feelings.” In *Martin Buber on Psychology and Psychotherapy*, edited by Judith Buber Agassi, New York, NY: Syracuse University, 1960, 129–138.

Dostoyevsky, Fyodor. *Notes from the Underground*. Ewing, NJ: Princeton University, 2004.

Foucault, Michel. “Nietzsche, Genealogy, History.” In *Language, Counter-Memory, Practice*, edited by D.F. Bouchard, Ithaca, NY: Cornell University, 1977, 139–165.

Kafka, Franz. *Letter to His Father*. New York, NY: Schocken Books, 1966.

———. *The Metamorphosis, In the Penal Colony, and Other Stories*. New York, NY: Schocken Books, 1995.

———. *The Castle*. New York, NY: Schocken Books, 1998.

Magny, Claude-Edmonde. *Les Sandales d’Empédocle*. Neuchatel: Éditions de la Baconnière, 1945.

Russell, Paul. “Strawson’s Way of Naturalizing Responsibility.” *Ethics* 102, 2: (1992) 287–302.

Strawson, P.F. *Freedom and Resentment and Other Essays*. London New York: Routledge, 2008.

Williams, Bernard. *Truth and Truthfulness: An Essay in Genealogy*. Ewing, NJ: Princeton University, 2004.