

"Responsibility" (Verantwortung)
 (for the *Handbuch der Politischen Philosophie und Sozialphilosophie*)
 by Joel Anderson

PENULTIMATE VERSION – NOT FOR QUOTATION OR CITATION!!!!

1. "Responsibility" (Verantwortung) is primarily a matter of being answerable for one's actions. Of the many senses of the term, the two distinct meanings that are most important within social and political philosophy can be illustrated in terms of two reasons a person might cite in denying that she was responsible for not saving someone from a burning building: that she was somehow prevented from doing so (she was paralyzed with fear or was injured) or that the duty to risk oneself in helping did not apply to her (but rather to the firefighters at the scene). In the first sense, saying that X is responsible for the consequences of an action (or decision not to act) emphasizes that X is related to the action in a way that makes the action X's own doing – and in such a way that attributions of blame or guilt (whether moral, legal, or otherwise) become appropriate. Typically, X is the agent that brought about the action and thereby its consequences. Call this "authorship-responsibility" [Verantwortung als Urheberschaft]. (It is close to what is sometimes called "retrospective responsibility.") In a second sense, to say that X is responsible for an action emphasizes the wrongness of failing to do something. This is sometimes referred to as "prospective responsibility"; but since this can also be attributed *post factum*, a better term is "obligation-responsibility" [Verantwortung als Verpflichtung].

There are other senses of "responsibility" as well, although they are more common outside of philosophical discussions. One has to do with physical causal relationships, as in "The faulty wiring was responsible for the fire." But this usage – Verantwortung als Ursache – is clearly parasitic on the core meaning as 'personal responsibility' (including collective persons). After all, we may blame the fire on the wiring, but we don't really *blame* the fire in the relevant sense. A further sense in which a person is sometimes said to be 'responsible' has to do with his or her having the character virtue of being particularly conscientious. Insofar as it is treated here, it falls under obligation-responsibility.

2. Historically, the term "responsibility" has been particularly prominent in the 20th century, but the roots of the various senses of the term go back to the beginning of Western philosophy. Two particular lines of historical development are worth highlighting: one focused on the notion of *ascribing* responsibility (and is related especially to the

sense of authorship-responsibility) (2.1); the other focuses on the notion of *taking* responsibility (with its ties to obligation-responsibility) (2.2).

2.1 Aristotle was deeply interested in the question of when it is just to hold someone responsible, and *two* features of his discussion have had a lasting impact on discussions of responsibility. First, he treats it as necessary for responsibility that the action is voluntary, in the sense that "the moving principle is in the agent himself" (Aristotle, *NE*, 1111a). An action is involuntary, according to Aristotle, if it is compelled (there is no alternative). Second, the agent must understand the effects of her action and the action is not voluntary if the results could not be foreseen. Both ideas have been controversial, but especially the requirement of freedom from compulsion has been problematized in debates over determinism and free will. For if every human action is unavoidably linked by deterministic laws of nature to a subsequent chain of consequences and a preceding chain of mental states and circumstances, there seems to be no place for the responsible agent as itself the "moving principle". Historically, there have been two main lines of response to this issue.

2.1.1 First, there are those who follow Aristotle in defending the idea that voluntariness is crucial for responsibility but insist that there is a way of understanding the responsible agent's relationship to her action such that the action can be said to be within her control or to genuinely come from within her in some other sense. Prominent figures here are Epicurus, Augustine, Aquinas, and Kant. Typically, the arguments rely on something non-physical – such as Reason or God -- playing the linking role. An important recent development along these lines has been the rise of "reasons-responsiveness" as a central requirement for the sort of "guidance-control" essential for responsibility (Fischer and Ravizza, 1998).

2.1.2 The second dominant approach is found among naturalistic philosophers who reject appeals to agency that lie outside the physical world -- but who still wish to retain (most of) our responsibility-related practices and institutions. Prominent figures include the Stoics, Hobbes, Hume, utilitarians, and, more recently, Daniel Dennett and various neurophilosophers. These theorists render responsibility and determinism compatible by loosening (or breaking) Aristotle's tight link between the conditions of being the agentic cause of the action and the conditions of ascribing responsibility (and with it, blame and punishment, praise and reward). These naturalist compatibilists see the grounds for holding agents responsible as lying not in the degree of causal or rational control the agent has but rather in considerations such as certain criteria being socially useful, evolutionarily

adaptive, or presupposed by important moral, political, or legal norms.

In this connection, three contemporary theorists have been particularly influential, although in very different ways. First, H.L.A. Hart (1968) developed a legal positivist conception of responsibility that shifted the focus from meta-ethical or metaphysical foundations to the "general justifying aim" behind a system of law, including attributions of responsibility. Second, P.F. Strawson (1963; see also Wallace 1994), argued that the question of whether it is possible to hold people responsible for many of their actions is not so much a question about the metaphysical causal relations as a question about presuppositions of interpersonal relations: without such a possibility, certain profound and essential features of the social world cannot be made sense of. Third, Harry Frankfurt (1988) argued that what matters for responsibility are not the antecedent causes or conditions of choosing, but the agent's attitude toward the actions; according to Frankfurt, even if we do not have alternate possibilities, we can be responsible for an action that we endorse in our "second-order desires" that the desires we have be effective in action.

2.2. A quite different tradition, centered on the importance of *taking* responsibility, has at least three distinct strands: civic republicanism, Augustinian Christianity, and existentialism. Within civic republicanism (e.g. Cicero and Hegel), we find the idea that individuals have not only general duties but also particular responsibilities as persons who occupy certain social and institutional roles, *for* certain persons or tasks, and answerable *to* specific authorities. The governor of a province, for example, might be responsible to the emperor for the defense of certain fortresses but to the merchants' guild for the maintenance of certain bridges. Within various monotheistic traditions (most influentially, traditions stemming from St. Augustine's *Confessions*), these obligation-responsibilities entail that one is not just *wrong* in failing to fulfill one's responsibilities but that one must also *answer* to God for such failings. This eschatological and relational notion of a final *reckoning* finds resonance in two ways in which contemporary understandings of responsibility has been influenced by existentialism's claim that the significance of events is not found in objective facts but rather determined by the attitudes that one takes toward them. First, there is the idea of 'Selbstwahl' (Kierkegaard, Nietzsche, Sartre, and perhaps Heidegger), as a matter accepting responsibility for one's life-history as a whole, rather than presenting oneself as a victim of circumstances. Second, there is the idea that one can always come to be confronted by others, even for what one reasonably thought, at the time, was the right thing to do. We must, on this view,

accept that the best of intentions are not enough to guarantee a clean conscience. This departure from a "Gesinnungsethik" (Max Weber's term for moral theories in which rightness is determined entirely by the rightness of intentions) in favor of a "Verantwortungsethik" (according to which moral agents must consider actual consequences as well) gained support and urgency in the wake of the atrocities of World War II as well as growing threats to the environment. Thus, in place of this more restricted scope of responsibility, thinkers such as Jaspers, Sartre, Jonas, and (most radically) Levinas have argued that the fact that the consequences and circumstances of our actions are not at our disposal does not relieve us of the obligation to courageously face the fact of our (perhaps inadvertent) involvement in evil and to take responsibility for the consequences of our actions.

3. Systematically, one can distinguish several strands within each of the two broader senses of "authorship responsibility" and "obligation responsibility."

3.1 In the current discussions about "authorship responsibility" five factors (and issues of ongoing controversy) stand out: the requirement of freedom (3.1.1), the role of attitudes of endorsing the motives for one's actions (3.1.2), the capacity to comprehend what one is doing (3.1.3), the scope of collective responsibility (3.1.4), and the range of appropriate responses to responsible wrongdoing (3.1.5).

3.1.1 One of the clearest grounds for saying that someone is not responsible for an action is that he was forced to do it – that he was not free. In many cases, it is obvious when the requirement of "negative" liberty (freedom *from*) is violated, but often the criteria are disputable, particularly in cases of threats. Is a soldier no longer responsible for violating human rights, for example, if she was threatened with dismissal if she participate? What if she was threatened with summary execution? These issues of judgment are further complicated by two factors: First, internal psychological compulsion (e.g., kleptomania) is often especially difficult to distinguish from an agent's deepest desires. Second, there are requirements of positive freedom – that one has the wherewithal to avoid a particular action – which are difficult to formulate in a way that avoids excusing all wrongdoing.

What these discussions remind us of, however, is that responsibility is closely connected to effective control by the agent. This leads to extensive, often highly technical discussions of the metaphysics of free will and determinism (for an overview, see Kane 2005), which cannot concern us here. But the key intuition is decidedly relevant for social and political philosophy: the idea that one could have done

otherwise, usually formulated as the "principle of alternate possibilities." Among other things, this principle provides a way of operationalizing the intuitive requirement that agents can be held responsible only for actions over which they have effective control. This principle has been debated extensively in recent decades, especially in response to Harry Frankfurt's argument that one can be responsible even if one couldn't have acted differently (whether in the strong metaphysical sense of impossibility or in the more practical sense of being coerced or threatened), if it is the case the action is one that one *would* have performed anyway, even if one *did* have other options. According to this family of views, the fact that a soldier is threatened with execution if she doesn't slaughter civilians doesn't relieve her of responsibility if she was already intent on committing the atrocity anyway.

3.1.2. There has been a widespread shift in approaches to ascribing responsibility, away from a focus on the metaphysics of causal interference to a focus on agents' attitudes. In part, this is to provide a way of determining whether one's motives for acting are authentically one's own, especially given the fact that so much of what clearly autonomous agents care about is the product of a developmental, life-historical process that they cannot control all the way back (Christman 1991). In this way, for example, religious attitudes resulting from early childhood socialization might be genuinely one's own or an external, responsibility-reducing influence, depending on whether one endorsed or rejected the attitude (or the process generating it).

Agents' attitudes are also seen as fixing the range of action-consequences for which one is to be held responsible. Particularly within the ethics of medicine, war, and public policy, the "doctrine of double effect" has been used to argue (roughly) that an agent can be held responsible only for those effects of his action that he intends, not merely for everything that he foresees. On this view, for example, even if a government knows that there is a strong correlation between literacy rates and suicide, it still ought not to be held responsible, in adopting a literacy program, for the suicides. Similar arguments have been used in contexts of "collateral damage" from wartime bombing and life-shortening use of powerful pain medication. These are subject to extensive controversy, not least of which because it is very difficult to determine the exact intentions of agents (or even oneself).

3.1.3. Psychological factors in responsibility also involve the requirement that one understand what one is doing. If one gives someone a drink that has been poisoned without one's knowledge, then one is not responsible for the death that was nonetheless brought about by one's action. Very sophisticated

discussions of this link between the agent's mind and a wrongful deed are to be found in the legal doctrine of *mens rea* (guilty mind). In addition to the highly contentious and widely varying principle that insanity and mental illness can make it impossible to hold someone responsible, legal theorists have developed several important distinctions between acting purposively, acting knowingly (aware of the wrongness of what one is doing but not intending it clearly and specifically, as in cases of "willful blindness"), and acting recklessly (where one is aware that one's actions endanger others, but doesn't care) (Hart 1968, ch. 6; see also *infra* 3.2.2). In philosophical contexts, much attention has been paid to the question of whether someone knows right from wrong, or in more recent versions, whether she meets the criterion of being (moderately, weakly, or strongly) reasons-responsiveness, as a necessary requirement for responsibility (Fischer and Ravizza 1998).

3.1.4. Is it only individuals who can be responsible? Many of the greatest moral outrages in history have resulted from the combined efforts of large numbers of people. It is straightforward enough to hold leaders responsible for commanding or encouraging others to commit atrocities, but conceptual complications quickly arise, particularly as atrocities have increasingly been caused by anonymous, amorphous, or distributed processes. What does it mean, for example, to say that "Germans bear responsibility" for the Holocaust, or that Exxon is responsible for the Valdez oil spill? Reductionistic individualist approaches claim that responsibility is fundamentally an individual concept and that talk of collective responsibility involves shorthand aggregation of an individual phenomenon. Defenders of collective, corporate, shared, or joint responsibility argue in different ways that, under certain conditions, a collective can be held responsible for actions, either because they form a joint agent of the action or because they stand in relations to wrongdoers in a way they distributes responsibility to them (May 1992). Note, however, that sharing in responsibility need not reduce the portion that each person bears.

3.1.5. There is also disagreement about what appropriately follows from demonstrating that someone is responsible for an action (good or bad). Some argue for a very tight connection between ascription of responsibility and what follows from that, seeing them as two aspects of the same social status. And especially for them, there are concerns that the acknowledging the possibility of collective responsibility seems to entail mass punishments for nations, groups, etc. But there is also room for distinctions among what responses might be entailed by the ascription of responsibility. For example, it might not require punishment but rather an obligation to acknowledge guilt, to express remorse, to atone

for what one did, or to provide restitution (Jaspers 1946). One prominent example of such alternatives, are "truth and reconciliation" commissions (as in South Africa), where concerns with punishment are suspended so as to encourage wrongdoers to accept responsibility for their deeds.

With regard to praiseworthy actions, the question of what follows from attributing responsibility focuses on discussions of "merit" and "luck". In recent discussions of **social justice** and **equality**, it is hotly debated whether a fair distribution of social goods requires minimizing the role of luck (e.g., happening to be born with special talents or affluent parents), such that they are accorded as much as possible on the basis of "merit", that is, on the basis of factors for which one is responsible, a position termed "luck egalitarianism" (discussed in Hurley 2003).

3.2. Obligation responsibility [Verantwortung als Verpflichtung]

As was noted at the outset, talk of responsibility is also used to highlight *duties* that agents have, for which others hold them accountable. In this sense, saying that some lacks responsibility is not to *excuse* her but precisely to *criticize* her. Central here are the notions of responsibilities as duties (3.2.1), negligence and responsibility for omissions (3.2.2), self-regarding negligence (3.2.3), and the willingness to take responsibility beyond the limits of reasonable ascriptions of responsibility (3.2.4).

3.2.1. The most straightforward sense in which one can be criticized for not fulfilling one's responsibilities is when one has specific obligations to look after something or someone, often as a function one's role: parents are responsible for their children's welfare; lifeguards are responsible for swimmer's safety; etc. The notion of role-responsibilities (e.g., a lifeguard's) is especially important in professional ethics. Interestingly, a strongly role-based conception of responsibility can also serve to narrow the scope of others' duties and responsibilities: if there is a lifeguard on duty, saving a drowning child is "not my job."

3.2.2. Even when one does not have a specific obligation to do something, a failure to act can render one responsible for bad consequences. One's failure to put protective covers on the outlets in one apartment may render one responsible for the electrocution of a 2-year-old nephew. Here, it is not clear whether it makes most sense to conceptualize this as one's omission as case in which one is the agent [Urheber] of the death or rather that one is to be blamed for failing to meet one's obligations. But either way, as is so often the case in discussions of responsibility, the question of where

to draw the lines appears to be answerable only on the basis of normative considerations (such as the "reasonable person" standard) rather than metaphysical criteria. Note that in some legal contexts (typically civil cases), one can even be held responsible, on grounds of "strict liability," although one had no idea that one was doing something wrong or took extensive measure to minimize risk.

3.2.3 What one might call "self-regarding negligence" ["selbstbezogene Nachlässigkeit"] involves forms of irresponsibility in which agents fail to take appropriate steps to ensure that they are no longer able to act responsibly. If someone knows that she has a tendency towards violence when drunk but goes ahead and gets drunk, or if one is aware of a tendency for epileptic spasms but fails to take the medicine, then she is responsible for ensuring that these otherwise excusing conditions do not obtain. Letting them occur undoes the exculpating force of the condition. ("Culpable ignorance" is a similar case, in which one is negligent in failing to acquire information, such as learning the laws of the land.)

One implication of this point is that a failure to modify one's character might be irresponsible. As Dennett puts the related point: "Doesn't a considerable part of being a responsible person consist in making oneself unable to do the things one would be blamed for doing if one did them?" (1984: 135). There are important questions here as to the extent to which agents can be obligated to take proactive steps to reduce the likelihood of acting irresponsibly in the future. And this might seem to create sense of "responsibility for self" (Ch. Taylor) that has strong perfectionist implications. Indeed, as more and more therapies become available for conditions that have until now been excusing (e.g., anti-psychotic medication), questions about the limits of negligence are likely to become more pressing.

3.2.4. At the outer limits of responsibility, profound questions arise as to whether there isn't something cowardly in hiding behind excusing conditions, even when they are true. Once it becomes clear that every action is a response to a situation produced by a sequence of events over which one never has ultimate control, "being responsible" comes to have an ineliminable existential and normative dimension, given the underdetermination from third-person points of view.

In this sense, we may sometimes be obligated to "take responsibility" for consequences of actions that that were not performed intentionally or freely. At the same time, the eagerness of many politicians pronounce that they "take full responsibility" for unfortunate events that they could not

have prevented points to the danger of turning this into an empty gesture.

In any event, the widespread expectation that people do take responsibility, in some sense, for bad occurrences in which they are somehow involved highlights the sense in which responsibility is not simply a personal phenomenon (having to do with one's own moral conscience) nor a metaphysical phenomenon (having to do with the objective relations between the will, an action, and its effects) but also a social, intersubjective phenomenon, having to do with the maintenance of the mutual attitudes that partly constitute responsible human agency.

4. Bibliography

Aristotle, *Nicomachean Ethics*, Books II and III.

Bieri, P., 2001, *Das Handwerk der Freiheit*, München/Wien: Carl Hanser.

Christman, J., 1991, "Autonomy and Personal History," *Canadian Journal of Philosophy* 21:1.

Dennett, D., 1994, *Ellenbogenfreiheit*, Beltz Athenäum (*Elbow Room*, 1984).

Fischer, J.M. and M. Ravizza, 1998, *Responsibility and Control: A Theory of Moral Responsibility*, Cambridge: Cambridge University Press.

Fischer, J.M., 1999, "Recent Work on Moral Responsibility," *Ethics* 110:1.

Frankfurt, H., 1988, *The Importance of What We Care About*. Cambridge: Cambridge University Press.

Hart, H.L.A., 1968, *Punishment and Responsibility*, Oxford: Oxford University Press.

Hurley, S., 2003, *Justice, Luck and Knowledge*, Cambridge: Harvard University Press.

Jaspers, K., 1946, *Die Schuldfrage*, Heidelberg: Schnieder.

Jonas, H., 1979, *Der Prinzip Verantwortung*, Frankfurt: Suhrkamp.

Kane, R., 2005, *A Contemporary Introduction to Free Will*, New York: Oxford University Press.

May, L., 1992, *Sharing Responsibility*, Chicago: University of Chicago Press.

Pothast, U. (Hg.), 1988, *Seminar: Freies Handeln und Determinismus* (2. Auflage), Frankfurt am Main: Suhrkamp.

Strawson, P., 1963, "Freedom and Resentment," in Watson (Hg.), 1982.

Wallace, R.J., 1996, *Responsibility and the Moral Sentiments*, Cambridge, Mass.: Harvard University Press.

Watson, G. (Hg.), 1982, *Free Will*, Oxford: Oxford University Press.

Weber, M., 1919/1988, "Politik als Beruf," in *Gesammelte politische Schriften*. Tübingen: Mohr.