Moral Responsibility

Forthcoming in the launch issue of *Oxford Bibliographies Online: Philosophy*

Introduction

Moral responsibility relates to many significant topics in ethics and metaphysics, such as the content and scope of moral obligations, the nature of human agency, and the structure of human interaction. This article focuses on compatibilist approaches to moral responsibility – that is, approaches that see moral responsibility as compatible with the causal well-orderedness of the world. This is partly because they have more to say about the nature of moral responsibility and the practices associated with it, and also because there will be a future bibliography on free will. The article also focuses mainly on the debates considered most significant by contemporary analytic philosophers. However, it also points to some earlier contributions and to some significant contributions from outside those debates. In particular, it is interesting that contemporary debates often focus on the agency of the responsible person, without attending to the forms of interaction in which that person may participate. However, as Peter Strawson pointed out in a seminal essay (see *Responsibility and the Reactive Sentiments*), moral responsibility is intimately related to our reactions to one another. Should those reactions be understood by reference to features of the person held responsible, or by reference to the relationship between persons where some action or

outcome is at issue, or even by reference to wider social and political structures? Moral responsibility also borders on a number of topics of great practical importance. These include responsibility under the law, the responsibilities of groups and organisations, accountability within organisations, and how distributive justice and individual responsibility are related. Again, this entry focuses largely on individual moral responsibility and only mentions a few social and legal discussions of responsibility with especial implications for how we think about individual responsibility.

General Overviews

A number of recent overviews give useful introductions, but offer different approaches to the topic. Duff (1998) and Eshleman (2001) are relatively accessible introductions; McKenna (2004) is more demanding and technical; Kutz (2002) is the most intellectually penetrating, but is orientated by concerns in law and jurisprudence rather than morality.

Duff, A. (1998) 'Responsibility,' *Routledge Encyclopedia of Philosophy*, ed. E. Craig, (Routledge, London, 10 vols.)

> [Useful short outline of the issues. See also the encyclopedia's entries on: Praise and Blame; Determinism and indeterminism; Free will.]

Eshleman, A. (2001/4) 'Moral Responsibility,' *Stanford Encyclopedia of Philosophy* [plato.stanford.edu/entries/moral-responsibility]

> [A useful short overview that gives some historical background, and locates current approaches in the light of Peter Strawson's influential contribution.]

McKenna, M. (2004) 'Compatibilism,' *Stanford Encyclopedia of Philosophy* [plato.stanford.edu/entries/compatibilism]

> [Lengthy, technical and proficient overview of compatibilist approaches to moral responsibility, with a further appendix on the most recent debates.]

Kutz, C. (2002) 'Responsibility' in J. Coleman & S. Schapiro (eds) *Oxford Handbook of Jurisprudence and Philosophy of Law*

> [Oriented toward legal debates. Nonetheless, a significant contribution arguing that the *relational* aspects of responsibility attribution are of critical importance. That is, we hold persons responsible within the context of particular relationships – personal, organisational or legal – and consider ourselves responsible to particular persons or bodies.]

Anthologies and Textbooks

There are several useful anthologies. Watson (1982) is still very useful, while the extensively revised second edition (2003) compasses more recent debates referred to in subsequent sections of this entry. Schoeman (1987) is an important collection of original papers. Fischer and Ravizza (1993) reprint many influential contributions – often these are cited in subsequent sections of this article. Honderich (n.d.) is also a

very useful on-line collection of influential papers. There is no textbook specifically on the topic of moral responsibility. However, Lucas (1993) and Matravers (2007) offer interesting, non-technical discussions touching on several different facets of responsibility.

Watson, G. (ed) (1982) *Free Will* (Oxford University Press, Oxford) (2nd edition 2003)

[Anthology of previously published twentieth century treatments of moral responsibility as well as free will – a very useful introduction. The second edition is much revised to include many more recent articles.]

Schoeman, F. (ed) (1987) *Responsibility, Character and the Emotions* (Cambridge University Press, Cambridge)

[A collection of original papers of very high quality.]

Fischer, J.M. & Ravizza, M. (eds) (1993) *Perspectives on Moral Responsibility* (Cornell University Press)

[An anthology that reprints many influential papers, from Strawson's 'Freedom and Resentment' onwards.]

Honderich, T. (ed) (n.d.) *The Determinism and Freedom Philosophy Website*, [www.ucl.ac.uk/~uctytho/dfwIntroIndex.htm]

[Comprehensive selection of influential articles, mainly from the last few decades but including a few extracts from classic texts. All are freely available on-line.]

Lucas, F.R. (1993) *Responsibility* (Clarendon Press, Oxford)

[Not strictly an overview nor a textbook, but succinctly treats many different aspects of responsibility, some of which are generally overlooked in contemporary discussions.]

Matravers, M. (2007) *Responsibility and Justice* (Polity, Cambridge)

[Begins with a thorough but accessible discussion of contemporary theoretical debates regarding moral responsibility. Though compatibilist in approach, Matravers argues that compatibilists tend to be too sanguine that all aspects of our practices of responsibility can be justified. Investigates both questions of responsibility in distributive justice (the debate about what has come to be called 'luck egalitarianism') and retributive justice (punishment).]

Free Will

As most readers will know, moral responsibility has often been discussed in connection with the question of free will. This generates the debate between incompatibilists (who believe that responsibility depends on a notion of freedom incompatible with natural causality, whether or not such freedom may actually exist) and compatibilists (who think moral responsibility is compatible with a naturalistic understanding of the world). A future entry will be devoted specifically to the topic of free will, so this section gives

only a few introductory pointers. Strawson (1994) provides a useful recent statement of a key argument for incompatibilism. Kane (2002) offers an advanced, authoritative overview of analytic approaches. Vihvelin (2003), Clarke (2000) and O'Connor (2002) are readily available and reliable encyclopedia entries. Finally, Levy and McKenna (2008) provide an up-to-date overview that requires familiarity with earlier debates.

Strawson, G., (1994) 'The Impossibility of Moral Responsibility,' *Philosophical Studies*, 75, 5-24

> [Trenchant restatement of the idea that moral responsibility depends on a notion of free will that is unavailable given our naturalistic worldview.]

Kane, R. (ed) (2002) *The Oxford Handbook of Free Will* (Oxford University Press, New York)

> [A thorough, advanced collection of overview articles by some of the most influential contemporary authors. It largely focuses on incompatibilist approaches but also includes a number of useful articles covering compatibilist contributions.]

Vihvelin, K. (2003/7) 'Arguments for Incompatibilism' *Stanford Encyclopedia of Philosophy* [plato.stanford.edu/entries/incompatibilism-arguments]

> [First of three Stanford entries, that together provide a comprehensive guide to the literature. This entry surveys the motivations for adopting an incompatibilist view. As well as arguments based on intuition, it may be argued that determinism prevents us from controlling our actions in the right sort of way for moral responsibility, or that determinism deprives us of the ability to do otherwise (this being, so it is held, essential to moral responsibility).]

Clarke, R. (2000/8) 'Incompatibilist (Nondeterministic) Theories of Free Will' *Stanford Encyclopedia of Philosophy* [plato.stanford.edu/entries/incompatibilism-theories]

> [This article surveys the three principal forms that incompatibilist accounts may take so far as the causation of actions is concerned: those that hinge on an absence of causality, those that trace responsibility back to an uncaused event, and those that invoke a specific sort of 'agent-causation.']

O'Connor, T. (2002/5) 'Free Will' *Stanford Encyclopedia of Philosophy* [plato.stanford.edu/entries/freewill]

> [This article considers the variety of things that philosophers have meant in speaking of free will: that we have the power to choose rationally, that we own our choices and actions (see *Identifying with our Actions*), or that we cause or control our actions. It also surveys some theological questions that originally prompted concerns about free will and responsibility.]

Levy, N. & McKenna, M. (2008) 'Recent Work on Free Will and Moral Responsibility,' *Philosophy Compass* 4:1, 96-133

> [Thorough, technical overview of debates within the last decade.]

Classic Texts

This entry focuses mainly on contemporary approaches to moral responsibility and their immediate precursors, but many authors in the history of philosophy have addressed this topic. Setting aside discussions animated by theological concerns – such as problems posed by divine foreknowledge of how we will act – the most important contributions are arguably those of Aristotle, Hume and Kant. Aristotle's discussion of excusing conditions in the *Nicomachean Ethics* [2000] is justly famous. Broadie (1991) and Meyer (1993) show how this brief analysis connects with Aristotle's wider theory of moral agency. Hume's moral theory [1748/51] is explicitly based in the qualities of character that elicit our esteem and reproach, and is an important inspiration for Strawson's influential account (see *Responsibility and the Reactive Sentiments*). Russell (1995) shows the sophistication of Hume's account and its place in his wider moral theory. Kant is often taken as the father of modern incompatibilist views about responsibility, because his view of moral obligation and moral character rests on the 'noumenal' freedom of the rational agent – a freedom that can never be known from our sensory or scientific experience of the world. Kant's account is deep and controversial, however, and he hardly discusses responsibility except in legal terms. His characteristic concern, explored at most depth in *Religion within the Boundaries of Mere Reason* [1998], is with how we should each regard all our actions as imputable to us, as reflecting an underlying and freely chosen moral disposition; but it is not clear that this provides any straightforward rationale for practices of blame and responsibility. Korsgaard (1996) draws on Kant's views to explore these practices and suggests that he may even be read in a broadly compatibilist fashion. Hill (2002) discusses Kant's account of punishment and shows that it does not invoke the metaphysics of free will. (See also Ripstein (1999) and Ripstein (2004), *The Value of Choice and Contractualist Developments*.)

Aristotle [2000], *Nicomachean Ethics,* trans. Roger Crisp (Cambridge University Press, Cambridge)

> [Book III §§1-5 investigate the conditions where blame is appropriate and the extent to which factors such as coercion, force of circumstances, and factual ignorance might excuse. As Broadie (1991) and Meyer (1993) show, however, these sections need to be read against Aristotle's broader picture of moral agency.]

Broadie, S. (1991) *Ethics with Aristotle* (Oxford University Press, New York)

> [In Ch. 3, 'The Voluntary,' Broadie argues that is it is a mistake to read Aristotle as concerned with questions of determinism. His concern is with the distinctive power of adult human beings to act in accordance with virtue. To develop this power and guard against its corruption, we must recognise and judge vicious persons and less than virtuous conduct, as well as their opposites.]

Meyer, S.S. (1993) *Aristotle on Moral Responsibility: Character and Cause* (Blackwell, Oxford)

> [Subtle analysis of Aristotle's account of moral agency. Argues that Aristotle's account does not invoke – as some readers of *NE* III.5 have thought – a responsibility for character that would be at odds with his emphasis on training

and habituation. Rather, it rests on the view that only moral agents can be the non-accidental, efficient causes of morally significant outcomes.]

Hume, D. [1748/51] *Enquiries concerning Human Understanding and concerning the Principles of Morals* (various editions)

[In the *Enquiry concerning Human Understanding*, 'On Liberty and Necessity' emphasises the regularity of human conduct, and that qualities of character are ascribed – whether in blame or praise – on the basis of this stability. The final Appendix to the *Enquiry concerning the Principles of Morals*, 'Of some verbal disputes,' rejects any clear line between the virtues, if they are understood as voluntary, and the talents: both are objects of praise, while their opposites attract censure and disapproval.]

Russell, P. (1995) *Freedom and Moral Sentiment: Hume's Way of Naturalizing Responsibility* (Oxford University Press, New York)

[Shows that Hume's view is more sophisticated than the 'classic' compatibilist view, that uncoerced action is free while compelled action – as opposed to action underlain by causal forces in general – is unfree. For Hume, our moral sentiments respond to qualities of character evinced by actions; it is qualities of character, rather than any liberty of the will, that underlie practices of responsibility.]

Kant, I. [1998] *Religion within the Boundaries of Mere Reason*, trans. Allen Wood & George Di Giovanni (Cambridge University Press, Cambridge)

[Book I offers Kant's most extended discussion of our propensity to subordinate morality to inclination (if not in action then at the level of motivation), and why we must regard this propensity as imputable to us: 'For if the moral law commands that we *ought* to be better human beings now, it inescapably follows that we must be *capable* of being better human beings' (p. 94, 6:50).]

Korsgaard, C. (1996) 'Creating the Kingdom of Ends: Reciprocity and Responsibility in Personal Relations' in her *Creating the Kingdom of Ends* (Cambridge University Press, Cambridge)

[A sophisticated Kantian account of responsibility that goes beyond Kant's own concern with self-imputation. Argues that responsibility pertains not to 'theoretical' facts – about abilities to do otherwise, say – but rather to the practical perspective that we must take in interacting with others as free and equal rational agents.]

Hill, T.E. Jr. (2002) 'Wrong-doing, Desert and Punishment' in his *Human Welfare and Moral Worth* (Oxford University Press, New York)

[Shows that Kant's account of punishment is not the retributivist account often attributed to him. In particular, Kant's account is not based on the moral deserts of the free will, but rather on the need for a coercive authority to uphold people's rights to a freedom of action that is compatible with like freedom for all.]

Contemporary Debates and their Precursors

The following four sub-sections are based on four of the most influential articles that continue to act as focal points in present debates. These are: (I) Jack Smart's utilitarian account of 'Free Will, Praise, and Blame; (II) Peter Strawson's 'Freedom and Resentment'; (III) Harry Frankfurt's 'Freedom of the Will and the Concept of a Person'; and (IV) H.L.A. Hart's 'Legal Responsibility and Excuses.' Although there are important points of contact between the lines of thought set out in each of the original articles referred to, each makes a distinctive point and has inspired a distinctive set of contributions up to the present day. These four approaches do not exhaust the field, even within analytic moral philosophy, and the two following sections point to further sets of issues that have emerged, or combine concerns raised by the four approaches set out here. While the first-mentioned, the utilitarian approach, is now less discussed, it would be fair to say that all of the other topics continue to attract intense debate. That said, the basic suggestion that issues from these approaches when taken together – that a compatibilist account of moral responsibility should be based in people's capacity to appreciate moral considerations and to exchange moral reasons – may be considered a matter of consensus.

(I) Utilitarian Instrumentalism

The utilitarian position – often referred to as an 'economy of threats' – remains the position that contemporary moral philosophers love to hate. Smart (1961) offers an especially clear and succinct utilitarian account of praise and blame; his is also a very clear example of an *instrumentalist* approach to responsibility attribution: praise and blame are tools that we use in order to encourage persons to make greater contributions to overall utility. Dennett (1984) is a refreshingly brusque confrontation with many intuitions that seem to support incompatibilism, and also sketches a basically utilitarian account of moral responsibility. While most writers on responsibility now reject – largely following Strawson, 1962 (see *Responsibility and the Reactive Sentiments*) – the idea that practices of responsibility are essentially instruments to secure social benefits, Arneson (2003) is important in revisiting Smart's contribution to suggest that it makes a point of enduring value.

Smart, J.J.C. (1961) 'Free Will, Praise, and Blame,' *Mind* 70, 291-306

> [Contends that praise and blame are about (i) 'grading' people and their contributions, and (ii) influencing them to do better.]

Dennett, D. (1984) *Elbow Room: The Varieties of Free Will Worth Wanting* (Clarendon Press, Oxford)

> [Sprightly run-through the 'bugbears' that, as Dennett holds, illicitly support intuitions about free will. Offers a broadly rule-utilitarian account of responsibility: we maintain practices of responsibility in order to support ourselves and one another in acting well.]

Arneson, R. (2003) 'The Smart Theory Of Moral Responsibility And Desert' in *Desert and Justice,* ed. Serena Olsaretti (Oxford University Press, Oxford)

[While critical of Smart, Arneson points out that his account gets something important quite right: practices of responsibility involve our exercising *influence* on one another in the cause of better conduct and outcomes.]

(II) Responsibility and the Reactive Sentiments

Peter Strawson's famous essay 'Freedom and Resentment' (1962) resituated the free will debate by highlighting the importance of 'reactive attitudes' such as resentment or gratitude to 'ordinary inter-personal relationships.' Against the free will-ists' resort to 'panicky metaphysics,' Strawson urges that the ways in which we divide responsible persons and actions from non-responsible agents and behaviours do not depend on anything so mysterious as an absence of causal determination, but rather on whether we are engaged in normal interpersonal relations – marked by such activities as quarrelling, loving, reasoning and much else besides. Affective responses like resentment and gratitude mark those relations, as they do not our relations with animals or the insane. (As Strawson observes, our relations with children occupy a crucial in-between place.) Against utilitarians like Smart, Strawson points out that affective responses and practices of responsibility would not work in the way they do, if we saw them as mere tools of enforcement or encouragement. To respond to agents in such a calculating way would mark an 'objective' attitude, as opposed to seeing the person as a fellow 'member of the moral community.' Strawson concludes, 'these practices [of reaction and responsibility], and their reception, the reactions to them [such as remorse], really *are* expressions of our moral attitudes and not merely devices we calculatingly employ for regulative purposes.' Pincoffs (1988) and Murphy and Hampton (1988) develop parallel lines of thought without obvious debt to Strawson. Watson (1987) and Russell (1992) are two of many illuminating responses to Strawson, and others are collected in McKenna (2007). Wallace (1994) is the most well-known and systematic development of Strawson, and is also important among the approaches which are grouped together under *Moral Reasons and Moral Address*.

Strawson, P.F. (1962) 'Freedom and Resentment,' *Proceedings of the British Academy* 48, 1-25 (Variously reprinted including Watson, 1982/2003; Fischer & Ravizza, 1993; Honderich, n.d. – see *Anthologies and Textbooks*)

[The single most influential essay in the literature on moral responsibility. Can be read as a direct response to the utilitarian instrumentalism of Smart (1961) and others. Practices of responsibility are not tools by which we manipulate others; they are part of what it is to relate to others as members of a moral community.]

Pincoffs, E. (1988) 'The Practices of Responsibility-Ascription,' *Proceedings and Addresses of the American Philosophical Association*, 61:5, 823-839

[Thoughtful but little-noticed essay that follows Strawson in emphasising that practices of responsibility are a mode of responding to the infliction of harm. If we were to reject those practices – say because of philosophical doubts about responsibility – we would be bound to settle on some alternative set of practices to deal with harms caused: but there seem to be no plausible alternatives.]

Murphy, J. & Hampton, J. (1988) *Forgiveness and Mercy* (Cambridge University Press, Cambridge)

[Thoughtful exchange of views particularly concerned with our feelings of anger and resentment in response to wrong-doing, the reasons that should lead us to forgive, and where it may be appropriate to endorse retributive feelings. (Not a response to Strawson as such.)]

Watson, G. (1987) 'Responsibility and the Limits of Evil: Variations on a Strawsonian Theme' in Schoeman (ed) (1987), reprinted in Fischer and Ravizza (eds) (1993) and G. Watson, *Agency and answerability: selected essays* (Oxford University Press, Oxford, 2004)

[Asks whether, and why, some persons might be exempted from the Strawsonian moral community. Uses a particularly horrible, real-life example of a wrong-doer: given his abusive up-bringing, was he in a position to appreciate the wrongness of the murders he committed? Influentially concludes that 'the boundaries of moral responsibility are the boundaries of intelligible moral address.']

Russell, P. (1992) 'Strawson's Way of Naturalizing Responsibility,' *Ethics* 102, 287-302

[Criticises Strawson for arguing that the reactive sentiments are part of our 'nature,' and therefore do not require justification or permit reasoning about whether they are called for.]

McKenna, M. & Russell, P. (eds) (2007) *Free Will and Reactive Attitudes: Perspectives on P.F. Strawson's 'Freedom and Resentment'* (Ashgate, Aldershot)

[Anthology reprinting important papers responding to Strawson. Includes Watson (1987), Russell (1992), an extract from Wallace (1994), as well as essays by Susan Wolf, Galen Strawson and others.]

Wallace, R J (1994) *Responsibility and the Moral Sentiments* (Harvard University Press, Cambridge, MA)

[Highly regarded study indebted to Strawson. Asks when it is fair to hold someone responsible and thus expose them to "reactive" emotions or various penalties. Suggests that this is fair where the person has actually done wrong (ie, his conduct is not excused) and where person has ability to respond to moral reasons (the lack of which, claims Wallace, exempts a person from responsibility).]

(III) Identifying with our Actions

Several essays of Harry Frankfurt's have been highly influential in discussions of the agency of responsible persons, although Frankfurt says relatively little about responsibility as such. While utilitarians emphasise social benefits and Strawson draws our attention to interpersonal relations, Frankfurt characteristically (Frankfurt 1969, Frankfurt 1971, Frankfurt 1987) emphasises aspects of self-hood. One surprising omission from Frankfurt's work and most subsequent commentary on him is the question of the standards by which a person should govern her choices or desires: Susan Wolf's contribution (1987) is important in observing this aspect. Schroeder and Arpaly (1999) suggest we should not necessarily look for the authoritative and responsible self in a person's most self-conscious beliefs about what she should do – just as important, if

not more so, are our habits and dispositions to respond to others. This point is also taken up by many of the accounts concerned with what should be attributed to a person, as discussed in the section, *Arguments Against Control*. Buss and Overton (2002) is a major collection of original papers on Frankfurt's work, many of which take up his concerns with the structure of the self and responsible agency.

Frankfurt, H. (1971) 'Freedom of the Will and the Concept of a Person,' *Journal of Philosophy* 68, 5-20 (Reprinted in Frankfurt, *The Importance of What We Care About: Philosophical Essays* (Cambridge University Press, Cambridge, 1988) and Watson (ed), 1982 – see *Anthologies and Textbooks*)

> [Suggests that the will is free when we are able to endorse – will, or identify with – our everyday, first order volitions. Someone who is subject to a compelling addiction is free insofar as she endorses her will to take the drug, and – claims Frankfurt – also responsible, even though she is, at another level, compelled to take the drug.]

Frankfurt, H. (1969) 'Alternate Possibilities and Moral Responsibility,' *Journal of Philosophy* 66, 828-839 (Reprinted in Frankfurt, 1988)

> [Trails the idea that a person might be fated to perform a particular act (ie, there are no 'alternate possibilities') and yet nonetheless wills to do so – and is accordingly responsible, even in the absence of 'an ability to do otherwise.' Much discussed in subsequent debates about free will.]

Frankfurt, H. (1987) 'Identification and Wholeheartedness' in Schoeman, ed. (1987) (see *Anthologies and Textbooks*) (Reprinted in Frankfurt's *Necessity, Volition, and Love* (Cambridge University Press, Cambridge, 1999) and in Fischer (ed) 1993 – see *Anthologies and Textbooks*)

> [One of several later essays where Frankfurt develops the notion of identification. Here Frankfurt emphasises that a higher-order process of considering and rejecting values can lead to a wholehearted identification with them. So a person 'makes up her mind,' and thereby takes responsibility for it. Says little concerning implications for our holding persons responsible, however.]

Wolf, S. (1987) 'Sanity and the Metaphysics of Responsibility' in Schoeman, ed. (1987)

> [Points out that a person may wholeheartedly identify with quite evil first order desires, perhaps owing to a perverted upbringing – to the point where we would not consider him responsible. Argues that a responsible agent needs to be morally 'sane' – that is, able to know right from wrong, and hence to evaluate her character (or first order desires) in a reasonable way.]

Arpaly, N. & Schroeder, T. (1999) 'Praise, Blame and the Whole Self,' *Philosophical Studies* 93: 161–188

> [Points out that we may fail to endorse, as a matter of explicit belief, our good actions ('inverse akrasia'). Against Frankfurt's account, which seems to imply that this removes praiseworthiness or even responsibility, the authors suggest an

'integrated self' account of responsibility, which suggests that we are responsible for all actions that more-or-less reflect our character.]

Buss, S. & Overton, L. (eds) (2002) *Contours of Agency: Essays on Themes from Harry Frankfurt* (MIT Press, Cambridge MA)

[Important collection of essays on Frankfurt with replies by Frankfurt himself. Many essays bear on the hierarchical model of the self and its relation to questions of responsibility, including those by Susan Wolf, Richard Moran, T.M. Scanlon, David Velleman, Gary Watson, J.M. Fischer, and Eleanor Stump.]

(IV) The Value of Choice and Contractualist Developments

H.L.A. Hart's (1968) more-or-less rule-utilitarian account of excusing conditions in the law has been influential to contractualist accounts of responsibility, especially T.M. Scanlon's. Hart himself strongly distinguishes between moral and legal responsibility, but it is obvious that a similar analysis might be offered in the moral case: We excuse a person who lacked the ability to conform to a moral norm, because morality is a system of requirements that is intimately related to our capacities to voluntarily govern our conduct. Scanlon (1998) develops this thought: in different ways, blame and punishment uphold moral standards, while upholding the value of our capacities of choice; Wallace (2002) and Williams (2006) respond. Lenman (2006) offers a slightly different take on the contractualist idea that practices of responsibility belong to reasonable terms of cooperation. Ripstein (1999) and Ripstein (2004) explicitly considers legal responsibility but should also be read as contributions to a wider political theory of responsibility (albeit one that is properly described as Kantian rather than contractualist).

Hart, H.L.A. (1968) 'Legal Responsibility and Excuses' in his *Punishment and Responsibility* (Oxford University Press, New York)

[Hart points out that law is meant to regulate people's conduct by acting on their capacities of choice. Where a person lacks the ability to meaningfully choose – as in cases of insanity or duress, for example – it therefore makes no sense to apply the same legal penalties.

Hart, H.L.A. (1968) 'Postscript: Responsibility and Retribution' in his *Punishment and Responsibility*

[Still useful as a discussion of the many different ways in which the words 'responsible' and 'responsibility' are used.]

Scanlon, T.M. (1998) *What We Owe to Each Other*, ch. 6 (Harvard University Press, Cambridge MA)

[Attacks a simple account of retrospective responsibility in terms of choice ("the forfeiture view"), for a more sophisticated "value of choice" view. We hold people responsible, not because they deserve whatever foreseeably accrues to them from their choices, but because people's capacity for choice is so important. But choice is only meaningful within a more or less predictable framework of consequences and responses.]

Wallace, R.J. (2002) 'Scanlon's Contractualism,' *Ethics* 112, 429-470 – plus response by Scanlon in the same issue at 510-513

> [Interesting critique of Scanlon by the author of one of the most influential books in the field. Of especial interest is the weight Wallace puts on a person's capacity to respond to moral reasons, if she is properly to be considered responsible. By contrast, Scanlon emphasises the authority of moral requirements, even if a person's vices seem quite fixed.]

Williams, A. (2006) 'Liberty, Liability, and Contractualism,' in N. Holtug and K. Lippert Rasmussen (eds.), *Egalitarianism: New Essays on the Nature and Value of Equality* (Oxford University Press, Oxford)

> [Critical discussion of Scanlon's account in *What We Owe to Each Other*.]

Lenman, J. (2006) 'Compatibilism and Contractualism: The Possibility of Moral Responsibility,' *Ethics* 117, 7-31

> [Suggests that moral responsibility represents a good contractualist bet. Imagining that one did not know what sort of moral character one would have, what sort of social arrangements would one choose? Given that human beings have some ability to live up to moral norms, argues Lenman, arrangements based on practices of responsibility would be a fair bet.]

Ripstein, A. (1999) *Equality, Responsibility and the Law* (Cambridge University Press, Cambridge)

> [Important recent legal and political discussion. Disavows the "voluntarism" (the focus on individual capacities underlying responsible agency and the fairness of retrospective responsibility) of many moral and legal accounts of responsibility, by suggesting that practices of responsibility are intrinsic to maintaining fair terms of interaction.]

Ripstein, A. (2004) 'Justice and Responsibility,' *Canadian Journal of Law and Jurisprudence* XVII:2, 361-386

> [Focussing on the law but again in the light of a wider political account. Important in developing the author's 'reciprocity conception of responsibility, which supposes that responsibility must be understood in terms of norms governing what people are entitled to expect of each other.']

Moral Reasons and Moral Address

As noted under *Identifying with our Actions*, Frankfurt-style approaches do not tend to consider the question of what standards a responsible agent should follow, while the utilitarian picture tends to suggest applying social pressure rather than the exchange of reasons. By contrast, Strawsonians and contractualists emphasise how responsible persons to respond to one another and to right reason. This section points to a group of approaches that relate to a prominent idea in RJ Wallace's (1994) account: that it is fair to hold a person responsible to the extent that she has the ability to respond to moral reasons (see *Responsibility and the Reactive Sentiments*). Wolf (1990) is an equally

important statement of this idea. (See also Korsgaard (1996) under *Classic Texts* for a sophisticated Kantian statement of an allied view.) Darwall (2006) offers a recognisably Kantian account, developed more systematically as a theory of morality. These authors all emphasise that practices of responsibility involve addressing the wrong-doer with moral reasons. Watson (1987) is often cited as coining the term 'moral address,' and the term is made more explicit in McKenna (1998); Hampton (1984) already suggested a similar idea with regard to punishment. Fischer and Ravizza (1998) is an influential work that also stresses the moral competence of the responsible person, although practices of moral address and mutual accountability play little role in their account. Like Frankfurt, they are concerned with individual agency; Watson (2001) accordingly criticises their subjective notion of 'taking responsibility.'

Wolf, S. (1990) *Freedom within Reason* (Oxford University Press, Oxford)

[Attacks Frankfurt-style 'Real Self' views and free will based 'Autonomy' views in favour of a 'Reason' view. The virtuous person who cannot act badly is responsible; the mad, stunted or brainwashed person who cannot act well is not; the rest of us, who are less than fully virtuous, may be nonetheless 'normatively competent,' hence responsible for our deeds.]

Darwall, S. (2006) *The Second-Person Standpoint* (Harvard University Press, Cambridge MA)

[Account of morality based on each person's authority to address others – 'second persons' – with moral claims. Accountability is therefore basic to ethics: the standing to address claims is coeval with the authority to press claims when these have been breached. Argues that this presupposes a shared competence to determine and act on moral reasons.]

Watson, G. (1987) 'Responsibility and the Limits of Evil: Variations on a Strawsonian Theme' in Schoeman (ed) (1987) (see *Anthologies and Textbooks*) (Reprinted in Fischer and Ravizza (eds) (1993) and G. Watson, *Agency and Answerability: Selected Essays* (Oxford University Press, Oxford, 2004))

[Influential response to Strawson 1962 (see *Responsibility and the Reactive Sentiments*) that argues that 'the boundaries of moral responsibility are the boundaries of intelligible moral address.']

McKenna, M.S. (1998) 'The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism,' *The Journal of Ethics* 2, 123–142

[Discusses and supplements the relation between Strawson 1962 (see *Responsibility and the Reactive Sentiments*) and Watson (1987), focussing on the role of moral address in practices of responsibility.]

Hampton, J. (1984) 'The Moral Education Theory of Punishment,' *Philosophy and Public Affairs*, 13:3, 208-238

[State punishment expresses the public view of what ought to be prohibited, and educates anyone who does not yet appreciate this: 'Wrong occasions punishment not because pain deserves pain, but because evil deserves correction.' Related

views may be found in Feinberg (1970), Fingarette (1967), Fingarette (2004), (see *Early Contributions*) and Hampton & Murphy (1988) (see *Responsibility and the Reactive Sentiments*).]

Fischer, J.M. & Ravizza, M. (1998) *Responsibility and Control: A Theory of Moral Responsibility* (New York: Cambridge University Press)

> [Argues that responsibility relates to 'guidance control' over one's actions, which rests on (i) moderate responsiveness to reasons and (ii) taking responsibility for one's capacity to control one's actions.]

Watson, G. (2001) 'Reasons and Responsibility,' *Ethics* 111:1, 374-94, reprinted in Watson (2004)

> [Critical response to Fischer and Ravizza (1998), focusing on the notion of 'taking responsibility.']

Arguments Against Control

Many discussions of moral responsibility assume that we can specify what someone is responsible for in terms of what was under her control. This idea clearly relates to an important truth, but Merrihew-Adams (1985), and a number of more recent articles, have offered devastating objections to any simple claim that we are only responsible for what was under our control. This issue arises from at least two directions. First, the emphasis on control has often motivated incompatibilist intuitions. Sher (2005) points out that the control-condition is more often assumed than argued for; his 2006a argues that it conflicts with a surprisingly wide range of cases where we would normally consider a person responsible. Second, questions about control arise in connection with the approaches mentioned in the last section, that emphasise the responsible person's responsiveness to moral reasons. What should we say of the responsibility of the thoroughly selfish or obdurate person, whose vices seem so settled that it seems false to say that she can respond to a certain class of moral considerations? It would be odd to think that such a person were thereby exempted from blame, or that we should relate to her just as we do to someone who exhibits the opposing virtues. Sher (2006b) examines and modifies the Humean view, that we judge a person's actions insofar as they reflect settled character traits. Smith (2005) and Scanlon (2008) develop parallel points. Such arguments are sometimes referred to as 'attributionist' because they suggest that many wrongful acts – be they chosen or spontaneous, non-deliberate or acts of omission – reveal flaws in a person's moral commitments. These actions and flaws are then attributed to the person, even if she seems unable to appreciate that they are morally faulty or to alter her conduct for the better.

Merrihew-Adams, R. (1985) 'Involuntary Sins,' *Philosophical Review* 94:1, 3-31

> [Argues that we can be responsible for states of mind such as jealousy, hatred or ingratitude, and that these cannot, in many cases, be plausibly traced to voluntary choices or omissions. Moreover, to overcome one's tendencies to such states typically involves *taking* responsibility for them.]

Sher, G. (2005) 'Kantian Fairness,' *Philosophical Issues* (*Nous* annual supplement) 15, 179-192

> [Observes how little argument has been offered for the 'Kantian principle' that it is unfair to hold people responsible for acts/omissions over which they lack control. Argues that we necessarily hold people responsible from a perspective that is not the actor's own; our blame may therefore express moral demands that the actor fails to appreciate.]

Sher, G. (2006a) 'Out of Control,' *Ethics* 116, 285-301

> [Points out how often we hold people responsible for actions even though they could not help performing them: eg, distracted or neglectful omissions, panicky or ill-judged actions, actions based on lack of moral insight or appreciation.]

Sher, G. (2006b) *In Praise of Blame* (Oxford University Press, Oxford)

> [Advocates a modified Humean account of blame (see *Classic Texts*), whereby actions are blameworthy insofar as they reflect stable, non-virtuous character traits. Sher argues that we (rightly) blame people for actions and outcomes that stem from 'the interplay of the… desires, beliefs and dispositions that… make [the wrong-doer] the person he is.']

Smith, A. (2005) 'Responsibility for Attitudes: Activity and Passivity in Mental Life,' *Ethics* 115, 236-271

> [Argues that practices of responsibility attribution relate to 'judgment-sensitive attitudes' rather than choices as such. Thus we can be responsible for forgetful omissions, failures to notice morally salient facts, involuntary reactions such as grief (or lack of), etc. All these are revealing of our attitudes to others and our evaluative judgments, and hence legitimate bases of moral appraisal.]

Scanlon, T.M. (2008) *Moral Dimensions: Permissibility, Meaning, Blame* (Belknap, Cambridge MA)

> [Account of blame based on the thought that, in general, we blame others within the context of particular relationships. To blame is to express our sense of the moral demands involved in that relationship, and how it has been undermined by some form of wrong-doing. Such practices do not ride on the idea that the wrong-doer is necessarily able to control his wrong-doing.]

Earlier Contributions

This section points to a number of thoughtful contributions that do not, for the most part, attract attention in contemporary philosophical discussions, although some (such as Austin 1956-7, Feinberg 1970, Williams 1993 and Williams 1997) are often cited. Often this neglect owes to their failure to fit some recognised school, or perhaps to a less technical approach than is currently favoured. It may be, then, they can help us to appreciate aspects of moral responsibility that are insufficiently attended to in contemporary discussions. Joel Feinberg and Herbert Fingarette are both important philosophers who collected many of their essays into stimulating collections (Feinberg,

1970; Fingarette 1967, 2004). Austin (1956-7), McKeon (1957) and Ricoeur (1992) are all revealing essays by important philosophers who otherwise did not write on responsibility as such. Bernard Williams often criticised the modern concern with voluntariness and choice (cf *Arguments Against Control* above): two of his most explicit treatments of responsibility are Williams (1993) and Williams (1997), the former being in part a response to Adkins (1960).

Feinberg, J. (1970) *Doing and Deserving: Essays in the Theory of Responsibility* (Princeton: Princeton University Press)

>[A set of classic essays on action and responsibility, often inflected by legal concerns but usually with an eye to moral responsibility too.]

Fingarette, H. (1967) *On Responsibility* (New York: Basic Books)

>[A set of classic essays. 'Acceptance of Responsibility' anticipates many later discussions, such as Susan Wolf's, by taking the example of psychopathy and argues that responsibility attributions are intelligible only insofar as they connect up with a person's existing moral concern.]

Fingarette, H. (2004) *Mapping Responsibility* (Open Court, Chicago)

>[A collection of beautifully succinct essays, summarising a lifetime's careful reflection on many aspects of responsibility.]

Austin, J.L. (1956-7) 'A Plea for Excuses,' *Proceedings of the Aristotelian Society* 57, 1-30

>[A set of stimulating provocations rather than a theory or account of excuse. Probes the concept of action and how actions may go wrong through our ways of talking about this. (Reprinted in his *Philosophical Papers*, J.O. Urmson & G.J. Warnock, eds. New York: Oxford University Press, 1979.)]

McKeon, R. (1957) 'The Development and the Significance of the Concept of Responsibility,' *Revue Internationale de Philosophie*, XI, no. 39, 3-32

>[A historical study of the concept, stressing that the original motivations for taking up responsibility as a topic of theoretical analysis were practical and political – in particular, to avoid controverted questions of metaphysics and moral psychology. Considers JS Mill and Max Weber, among others.]

Ricoeur, P. (1992) 'The Concept of Responsibility: An Essay in Semantic Analysis' in his *The Just*, trans. D. Pellauer (University of Chicago Press, Chicago)

>[Demanding and rich essay. Analyses the concept both in relation to the fundamentals of human agency and in relation to contemporary concerns with risk, indemnification, and – following Jonas (1984) (see *Social and Political Approaches*) – responsibility to future generations.]

Williams, B. (1993) *Shame and Necessity* (University of California Press, Los Angeles)

[Argues that the ancient Greeks had a sophisticated account of responsibility attribution, which avoided the mistaken emphasis on voluntariness Williams finds in modern morality, or at any rate, in philosophical interpretations of this. Contends that there may be different valid conceptions of responsibility in different contexts, but all combine four elements: *cause* [of action or outcome], *intention* [of the actor], *state* [of the actor' mind] and *response* [which we deem morally required of the actor].]

Williams, B. (1997) 'Moral Responsibility and Political Freedom,' *Cambridge Law Journal* 56: 96-102, reprinted in his *Philosophy as a Humanistic Discipline* (ed. A.W. Moore, Princeton, NJ, Princeton University Press, 2008: 119-25)

[A succinct statement of Williams' view that (i) responsibility rests on the four elements specified in *Shame and Necessity*, (ii) political freedom rests on the state's punishing only voluntary acts, and (iii) the idea of 'voluntariness' should not be regarded as metaphysically 'deep.' The edition of the journal also includes a reply by Antony Duff and Andrew von Hirsch.]

Adkins, A.W.H. (1960) *Merit and Responsibility* (Clarendon Press, Oxford)

[Argues that the Greeks lacked modern, Kantian notions of duty and fairness in assigning responsibility. Important largely as a foil for Williams (1993).]

Social and Political Approaches

The social and political ramifications of responsibility are enormous, and not often hinted at in contemporary philosophical discussions of moral responsibility. The following represent some exceptions to this tendency, in addition to those included in the previous section. Barnes (2000) stands apart as a contribution from sociological theory, while the others – for all their diversity – pose political questions about divisions and attributions of responsibility that are neglected when one focuses on the agency of individual persons, or frames matters in terms of 'moral reasons' in the abstract. Jonas (1984) underlines modern human beings' collective responsibility with regard to the existence of future generations, and the responsibility of the 'statesman,' but does not consider the division of responsibilities between individuals. Vickers (1973), Vickers (1980), Williams (2006), and Williams (2008) are concerned with the institutional conditions needed for responsible conduct: how persons may obtain a concrete sense of the 'reasons' they should act on, and how they should hold one another accountable. Smiley (1992) highlights questions of power and belonging, while Bovens (1998) considers how individuals ought to be held accountable when organisations act badly, and how responsible conduct and responsible organisations can be encouraged.

Barnes, B. (2000) *Understanding Agency: Social Theory and Responsible Action* (Sage, London)

[Barnes' sociological background and frank relativism will deter most philosophers, but those who persist will find a wealth of insight and argument. Barnes' central thought is that to consider a person as 'responsible' is to accord a social status, that corresponds to effective rights to participate in social life.]

Jonas, H. (1984) *The Imperative of Responsibility* (University of Chicago Press, Chicago)

[Argues that modern humankind's new power to destroy nature creates a historically novel responsibility toward future generations.]

Vickers G. (1973) *Making Institutions Work* (Associated Business Programmes, London)

[Vickers' work fitted no recognised discipline, let alone any recognised school of philosophy. The essays in this book, and in the following entry, are an important contribution to understanding modern institutional structures, and in particular the contributions and qualities – the responsibilities and sense of responsibility – that they demand of their members.]

Vickers, G. (1980) *Responsibility: Its Sources and its Limits* (Intersystems, Seaside CA)

[A set of essays that are effectively a companion volume to the previous entry.]

Williams, G. (2006) '"Infrastructures of Responsibility": The Moral Tasks of Institutions,' *Journal of Applied Philosophy*, 23:2, 207-221

[This and the following article consider why responsibility is such an important quantity in liberal societies. This article argues that it is related to the ways in which modern societies systematically – though often faultily – divide responsibilities across persons, roles and institutions.]

Williams, G. (2008) 'Responsibility as a Virtue,' *Ethical Theory and Moral Practice* 11:4, 455-470

[Considers what is involved in praising someone as 'responsible,' relating this to contexts of action that at the same time impose many moral demands on us and enable us to fulfil these.]

Smiley, M. (1992) *Moral Responsibility and the Boundaries of Community: Power and Accountability from a Pragmatic Point of View* (University of Chicago Press, Chicago)

[Argues that judgments of blame and responsibility hinge, in ways that philosophical discussions do not usually acknowledge, on configurations of power, social roles, and the boundaries of community – that is, who we are expected to take into consideration, or to answer to.]

Bovens, M. (1998) *The Quest for Responsibility: Accountability and Citizenship in Complex Organizations* (Cambridge University Press, Cambridge)

[Investigates how regulation, organisational reform, and different modes and channels of accountability can address irresponsibility on the part of institutions.]