

## **RETHINKING RESPONSIBILITY: THE ROLE OF JUDGEMENT AND BELIEF**

There are two standard conditions of responsibility: the knowledge condition stating that an action can only be responsible to the extent that its agent knows what she is doing and the control condition stating that a person can be held responsible only for things that are up to her. This paper will aim to explore the roles of judgement and belief in rethinking these two conditions.

### **1. The control condition: replacing choice with judgment**

What alternative formulations of the control condition have in common is the basic thought that a person can be held responsible only for things that are up to her (cf. Radoilska 2013, pp. 3–31). One intuitive way of fleshing out this basic thought is provided by the notions of choice or discretion over what one is doing. The underlying ambition is twofold. Firstly, to eliminate obvious instances of coercion in which what one is doing cannot be meaningfully treated as the expression of one's agency. Secondly, to eliminate the so-called 'mere behaviours', such as twitches, irresistible urges or any other aspect of the workings of the body that, similarly, leave little room for personal agency. Once these inner and outer limitations and impediments of agency have been set aside, we are left with voluntary actions and omissions as primary objects of responsibility ascriptions. This is because they appear to be the only plausible candidates satisfying the control condition formulated in terms of choice and discretion: voluntary actions and omissions are up to the agent by being amenable to her direct conscious effort. On this view, other things could be eligible objects of responsibility ascriptions to the extent that they are consequence of prior voluntary actions or omissions. Such secondary objects of responsibility ascriptions include attitudes and dispositions. In other words, moral praise or blame would be an appropriate response to them in so far as they result, albeit indirectly, from a person's choice or series of choices. Aristotle's conception of responsibility for character is sometimes interpreted along these lines (e.g. Levy 2005). On this interpretation, cowardice, for instance, is properly blamed rather than pitied even though, once settled, the disposition is no longer under the agent's direct control. The thought is that, having chosen to act on inappropriate or excessive feelings of fear, time and again, she has also chosen to become a coward. As a consequence of many voluntary actions in the past, a blameworthy disposition, such as cowardice is something for which an agent is rightly held responsible.

Importantly, the derivative status of responsibility for dispositions and corresponding attitudes is supported by the same intuition that places voluntary actions and omissions at the heart of responsible agency: choice or discretion over what one is doing is the hallmark of the required sort of control as it offers an agent the chance to avoid blame viz. deserve praise, *if she wanted to*. Choice so understood may but does not have to include the ability to do otherwise, which following Harry Frankfurt's seminal critique (1969) is mostly discussed under the heading of Principle of Alternate Possibilities or PAP (cf. Widerker and McKenna 2003). While recent work in experimental psychology offers some evidence that most people tend to identify their own agency with an ability to do otherwise in line with PAP (Deery et al. 2013), philosophical accounts of responsibility in terms of voluntary control tend to prefer a less sanguine notion, close to what Fisher and Ravizza (1998) termed 'guidance' as opposed 'regulative' control. That is to say, the control condition would be satisfied, on such accounts, as long as the agent chooses to do what she is actually doing independently of whether she could have done otherwise. Although less demanding than regulative control, guidance control is able to maintain the underlying distinction between activity and passivity, between things that are up to a person in her capacity of agent and things that are merely happening to her.

R. Jay Wallace's account of responsibility in terms of reflective self-control offers a helpful example of this trend. According to Wallace, reflective self-control is best understood as the conjunction of two powers: firstly, the power to grasp and apply relevant, especially moral reasons; and secondly, the power to control or regulate one's behaviour by the light of such reasons (1994, p. 157). These two powers are meant to replace the ability to do otherwise with a kind of normative competence that would enable us to address the ethics of responsible agency without (undue) reliance on the metaphysics of free will (cf. Wallace 2006, Ch.7). Crucially, this normative competence of reflective self-control interprets the relationship between voluntary actions and omissions, on the one hand, and attitudes and dispositions, on the other, in a relevantly similar way to that grounded in the intuitive notions of choice and discretion sketched earlier. As Wallace (1994, p.131) points out:

Particular states of emotion or feeling, however, are not the sorts of states that can directly be controlled by the reasons expressed in moral principles: such states as love, esteem, and goodwill are generally not states that could be produced by the belief that there are moral considerations that make them obligatory. This is why we cannot plausibly interpret moral

obligations as governing the quality of peoples' will where such qualities are construed broadly, to encompass emotions and feelings quite generally.

Since attitudes are not subject to the agent's direct voluntary control or discretion, they are not considered as appropriate objects of responsibility assessments in their own right. Whenever such assessment becomes fitting, it is so by extension and as a result of the responsibility incurred for voluntary actions and omissions that are at the agent's immediate discretion.

Critics of the control condition in terms of choice and discretion often target, explicitly (e.g. Smith 2005, 2008) or implicitly (e.g. Hieronymi 2014), Wallace's account of reflective self-control as paradigmatic formulation of the mistake they aim to expose and correct. As Angela Smith summarises her proposed alternative in a recent paper (2008, p. 369):

...the particular sort of non-volitional view I seek to defend..., which I have elsewhere called the rational relations view, makes rational judgement rather than choice or voluntary control the base condition of moral responsibility. To say that an agent is morally responsible for something, on this view, is to say that that thing reflects her rational judgement in a way that makes it appropriate, in principle, to ask her to defend or justify it.

On this view, the advantage that judgment is expected to have over choice and discretion is twofold. Firstly, it helps remedy an unduly restrictive interpretation of what is up to an agent. Secondly, it provides a criterion of responsible agency that tracks better the variety of ordinary practices of holding each other responsible. These two aspects, broadening the scope of legitimate objects of responsibility assessment and rethinking the control condition of responsibility, are intimately related. The link becomes apparent as soon as we consider the set of phenomena that are given the prominence of paradigm cases on Smith's conception of responsibility. These phenomena include spontaneous attitudes, such as fear, envy, gratitude and admiration, on the one hand, and on the other, patterns of awareness, such as forgetting a close friend's birthday or being otherwise inattentive to other people's expectations and needs (Smith 2005, p. 236). In both cases, a moral reaction, such as blame or praise seems to be called for. Yet, neither spontaneous attitudes, nor patterns of awareness can be adopted or rejected at will. If so, the moral reaction to these two categories of phenomena has to be grounded in something else, an alternative to the agent's discretion or voluntary control. As Smith (2005, p.251) argues:

When we praise or criticize someone for an attitude it seems we are responding to something about the content of that attitude and not to facts about its origin in a person's prior voluntary choices, or to facts about its susceptibility to influence through a person's future voluntary choices. More specifically, it seems we are responding to certain judgments of the person which we take to be implicit in that attitude, judgments for which we consider her to be directly morally answerable. If this is correct, then it is a mistake to try to account for a person's responsibility for her own attitudes in terms of their connection to her prior or future voluntary choices, because that obscures the special nature of our relation to our own attitudes: we are not merely producers of our attitudes, or even guardians over them; we are, first and foremost, inhabitants of them. They are a direct reflection of what we judge to be of value, importance, or significance. I have suggested that it is in virtue of their rational connection to our evaluative judgments that they are the kinds of states for which reasons or justifications can appropriately be requested.

Importantly, judgment is not meant to provide a local criterion for responsibility only, serving as a counterpart of voluntary control or discretion with respect to attitudes and patterns of awareness, viz. dispositions. Instead, judgment is presented as a global criterion that accounts equally well for responsibility with respect to actions and omissions. In any instance, in which it is apt to allocate responsibility, the allocation is done on the same grounds, in response to a judgment that is attributed to a rational agent. So, unlike the interpretation of the control condition in terms of discretion and choice, the alternative based on judgement is not supposed to support a two-tier approach to responsibility but to treat both actions and attitudes as fitting objects of responsibility ascription in their own right.

Another noteworthy feature of the proposed alternative is that judgments do not need to be considered or endorsed by the agent in order to be deemed as rational and so fit to attract a reaction in terms of moral praise or blame. The following vignette proposed by Thomas Scanlon (2008, p. 195) offers a helpful illustration:

... a man who firmly rejects racist views but who nonetheless sometimes finds himself thinking, when he sees people of a different race, that their skin colour is a reason for regarding them as inferior and preferring not to associate with them. We may suppose that when such a thought occurs to him he is appalled by it and he rejects his thoughts as mistaken and shameful. But they continue to occur nonetheless. The fact that these reactions are contrary to his considered judgment – that he 'disowns them' – makes a significant difference to our assessment of this person. It changes the overall picture of what he is like. But it does

not erase the relevance of these attitudes altogether. They are still attributable to him, and their occurrence is still a moral defect.

Looking at the vignette, we can see why critics of the control condition should consider reflective endorsement or, as the case is, reflective rejection of one's attitudes as another object of responsibility assessment rather than a criterion of responsibility, albeit a secondary one. At the very least, reflection may be plausibly construed as a counterpart of voluntary control or discretion with respect to attitudes and patterns of awareness (cf. Levy 2014a, 2014b). What's more, some critiques of the control condition are effectively aimed at debunking the apparent significance of reflection for rational viz. responsible agency.

For instance, Pamela Hieronymi (2014) argues that undue emphasis on reflection is at the root of a common yet mistaken intuition, which links responsible agency with discretion and voluntariness. According to Hieronymi, this intuition takes the following form: 'whenever we control a thing, we do so by reflecting upon that thing, deciding how it should be, and then bringing about that it is that way' (p.5). As a result, intending is conceived as exercising control over the outcomes of one's actions in virtue of 'having in mind' possible alternatives (awareness aspect) and 'making up one's mind' in favour of one instead of another option (discretion aspect). Throughout the paper, Hieronymi stresses that the underlying model of responsibility in terms of control can be safely applied to voluntary actions. However, this is only because such actions, in so far as their performance is, by definition, at the agent's discretion, are rather peripheral to both exercise and assessment of rational agency. The kind of responsibility incurred for them is merely juridical, on a par with that a person would incur for the misbehaviour of her pet dog (p.10). Juridical responsibility is not self-standing. It derives from a more fundamental conception, which Hieronymi terms answerability. On this conception, 'to be responsible for something is to be open to certain sorts of assessments on account of that thing...and to be the appropriate target of certain sorts of reactions on account of it' (p.9). More specifically, 'one is answerable for one's intentional actions... just in case a request for one's reasons is given application' (p.12).

Hieronymi presents her conception of answerability as loosely connected to the idea of the question 'Why?' in a special sense that agents are in a position to answer in so far as their actions are intentional, which was introduced by Elizabeth Anscombe (1963). In a nutshell, the thought is that there is a series of related questions, including: 'Why are you  $\phi$ -ing?', 'What are you doing?' and 'What are you doing that for?' and the ability to answer such a

series of questions is the mark of intentional action: it is known by its agent in a special, unmediated kind of way, ‘without observation’. And so the underlying question ‘Why?’ in a special sense is denied application, inter alia, when the agent is not aware of her  $\phi$ -ing.

On Hieronymi’s conception, if the question ‘Why?’ in a special sense is given application, this is because the agent has settled the question whether to  $\phi$  for the reasons that made her action worth performing (p.14). This observation leads Hieronymi to conclude that, ultimately, answerability tracks an agent’s evaluative perspective on the world and her place in it. But such a stance is too fundamental to be left to a person’s discretion and so cannot be accounted for by a juridical conception of responsibility. As Hieronymi (p. 20) puts it,

...if we are, most fundamentally, responsible for our take on what is true or important or worth doing – if our take on these questions is the object of the assessments and reactions that are characteristic of holding someone responsible – then we cannot enjoy discretion with respect to those things for which we are most fundamentally responsible.

The upshot is a two-tier responsibility conception. The first tier is answerability or responsibility in the strict sense, which attaches to both actions and attitudes in so far as they reveal or embody an agent’s affirmative answer to the question of whether what she is doing is worth doing. The second tier is a lesser kind of responsibility, which includes elements of management and control over outcomes or items in the world that, as Hieronymi terms it, fall under a person’s jurisdiction. Answerability is relevantly similar to Smith’s rational-relations view. Both present the basic condition for responsibility ascriptions as a kind of awareness without discretion. Both anticipate a comeback of the control condition via reflection and, in order to forestall it, express the awareness requirement in terms of a first-order evaluative judgment. Yet, looking closely at the role conferred to judgment on these two accounts, it becomes apparent that it is expected to combine two tasks that stand uneasily together: to provide the underlying criterion for responsibility while at the same time being the main subject of responsibility assessments. To appreciate the difficulty arising, let us consider the implications that rejecting the control condition in terms of choice and discretion has for construing the other standard condition for responsibility, the knowledge or awareness condition.

## **2. The knowledge condition: replacing reflection with first-order belief**

To recap, at the start of the paper, the knowledge condition was introduced as stating that an action can only be responsible to the extent that its agent knows what she is doing. As with the control condition, this formulation gains intuitive support from thinking about the nature of intentional action. Unless a person knows what she is doing, what appears to be an action of hers would effectively amount to a mere behaviour. The ability to answer the Anscombian question ‘Why?’ that was briefly considered earlier is meant to track the special way, in which agents are aware of their own actions. Since this mode of knowing cannot be replicated by any observer, no matter how well-placed or informed, knowing what one is doing in this unmediated, i.e. non-inferential and non-observational way could plausibly be interpreted as the mark of intentional action. This is not to say that observers may never be better judges of what is been actually done than the agent herself. Instead, the thought is that the agent knowing what she is doing while doing it owes nothing of consequence to evidence or reflection. As Anscombe’s example of writing on a blackboard with one’s eyes shut indicates (1963, §29), monitoring one’s action may often improve performance: getting to see whether the things that I am writing turn out to be legible is handy as it allows me to control the quality of the outcomes. However, it does not affect my ability to answer the question ‘Why?’ that marks out my writing on the blackboard as being intentional. Checking on the outcomes is of no consequence in this respect.

Looking at this reconstruction, it becomes apparent why proponents of responsibility as sheer answerability would be tempted to appropriate an Anscombian perspective on the knowledge condition. Being unmediated, awareness of one’s intentional actions seems to satisfy both theoretical desiderata, which motivate answerability: independence from reflection and the confinement of discretion to the margins of agency. At the same time, however, the reconstruction offers a caveat as it reminds us of the drawbacks of applying directly insights from philosophy of action into ethical theory (and vice versa), one of the important contributions of Anscombe’s seminal work. As the extensive discussion of the so-called cistern example (§§23–26) compellingly shows, an agent’s intention is not a definitive guide for the responsibility she incurs in virtue of her intentional actions. The example is as follows:

A man is pumping water into the cistern which supplies the drinking water of a house. Someone has found a way of systematically contaminating the source with a deadly

cumulative poison whose effects are unnoticeable until they can no longer be cured. The house is regularly inhabited by a small number of party chiefs, with their immediate families, who are in control of a great state; they are engaged in exterminating the Jews and perhaps plan a world war. The man who has contaminated the source has calculated that if these people are destroyed some good men will get into power who will govern well, or even institute the Kingdom of Heaven on earth and secure a good life for all the people; and he has revealed the calculation, together with the fact about the poison, to the man who is pumping.

The example's complexity supports a series of descriptions, under which the action under consideration is intentional: by moving his arm up and down, the man is operating the pump; by operating the pump, the man is supplying water to the house etc. Each of these descriptions answers the question 'Why?' in a special sense. However, this question does not elucidate some ultimate reasons for action that only the agent is entitled to know. As Anscombe points out, 'earning a living' might be the description, under which the man would account for his pumping poison into the water system of the house in response to the question: 'Why are you moving your arm up and down?' Assuming that the reply is sincere, poisoning the inhabitants would still be what the man does in a way that engages his full responsibility. Although not intended directly, poisoning the inhabitants is an integral part of what the agent is doing in order to earn a living. The question 'Why?' in a special sense tracks intelligibility, not justification from the agent's perspective. Knowing what one is doing needed to answer this question does not include a normative component, such as believing that what one is doing is worthwhile in one way or another. Yet, both proponents and critics of the control condition discussed so far agree on defining the knowledge condition as comprising a distinct and irreducible normative component. This shared understanding of the knowledge condition is related directly to a second point of agreement between these otherwise competing accounts: responsibility assessments take the form of reactive attitudes shaping the boundaries and structure of the moral community. As the subsequent discussion will aim to show, endorsing these aspects undermines the initially plausible strategy of replacing choice and discretion with judgement, strengthening the knowledge condition in order to compensate for the rejected control condition. The success of this strategy hangs on whether it can offer convincing answers to the following questions: What makes a judgment that, to an observer, might seem embodied by actions of mine attributable to me in a sufficiently robust way in order to justify my being held answerable for that judgment? Shouldn't I know what my action is supposed to mean in order to be rightly held responsible for the expression of that meaning? So, in a way, the question about the

feasibility of the underlying approach to responsibility boils down to that of whether epistemic responsibility can be considered as more fundamental than, or at the root of, moral responsibility. With the control condition severely curtailed if not entirely abandoned, responsibility for actions is, fundamentally, modelled on responsibility for beliefs.

As stated earlier, a normative component of knowing what one is doing is recognised by both accounts that incorporate and accounts that reject the control condition of responsibility. For instance, the first power of reflective self-control, the normative competence that as we saw earlier defines what a person can be rightly held responsible for on Wallace's account, is about grasping and applying to one's situation relevant normative and in particular moral reasons. Similarly, evaluative judgment embodied in one's actions and attitudes is the mark of rational viz. responsible agency in both Smith's and, *mutatis mutandis*, Hieronymi's accounts. In other words, answerability is about morally relevant beliefs that a particular action effectively asserts, with or without its agent's reflective endorsement and, in some cases, without her being aware of having made that assertion. To return to Smith's original example, forgetting a close friend's birthday is a fitting target of reproach as it exhibits a lack of concern akin to the belief 'it is o.k. to put myself first' that obviously flouts the norms of friendship. Note that, clearly, the agent doesn't mean to express the objectionable belief for which he is resented: he does not forget his friend's birthday on purpose. This aspect of the vignette is significant. It is supposed to show that discretionary control over what one is doing is inessential for moral responsibility. However, the question arises of whether the agent is sufficiently aware of what he is doing so that the accepted knowledge condition of responsibility is satisfied, independently of the rejected control condition.

Smith and other proponents of the strategy under consideration (e.g. Arpaly 2003, Talbert 2013) believe that this concern can be addressed by adopting a form of internalism about reasons for action that was first outlined by Bernard Williams (1981). Central to this kind of internalism is the idea of an asymmetry between internal reasons, which are best expressed in the form of 'A has a reason to  $\phi$ ' and external reasons, which conversely are best expressed in the form of 'There is a reason for A to  $\phi$ '. Only the former category of reasons are reasons for action proper since they are either already part of what Williams termed the agent's subjective motivational set or can be reached from premises contained within this set, if the agent were to follow a 'sound deliberative route'. By contrast, the latter category effectively expresses third parties' expectations from or exhortations directed toward an agent. Such

external reasons may indeed point to internal reasons or reasons for action proper in so far as they are able to make a person acknowledge elements of her motivational set, to which she has not previously given due consideration. In other words, whenever an apparently external reason manages to move a person, this is only by the force of some of her own pre-existing though perhaps not fully realised internal reasons.

This picture is attractive to critics of the control condition of responsibility as it seems to offer a way out of the difficulty identified earlier: satisfying the knowledge condition for responsibility in the absence of discretion. Of particular interest is the perceived possibility of acknowledging as internal some reasons for action, of which an agent is reflectively unaware, including at the time of action, as long as these reasons can be plausibly derived from her existing motivational set. Following this line of reasoning, it becomes plausible to conceive of responsible actions as assertions of an agent's morally relevant beliefs, without the need for her to arbitrate which of these elements of her motivational set gets asserted on a particular occasion. While independent from the agent's second-order reflection, such actions are deemed as responsive to reasons of the agent's own, first-order evaluative beliefs which she holds and for which responsibility as answerability is incurred. The link between actions as assertions and the first-order evaluative beliefs that they assert is expected to fulfil two functions. The first is to give application to the Anscombian question 'Why?' showing that answerable actions satisfy the knowledge condition of responsibility without recourse to reflection. The second is to establish the agent's evaluative stance as ultimate object of responsibility assessments illuminating a distinctly normative core tracked by the knowledge condition.

The revised knowledge condition is faced with immediate difficulties as soon as it is applied to account for specific aspects of the practice of holding each other responsible, in light of the second central feature of the underlying strategy, according to which responsibility assessments take the form of reactive attitudes shaping the boundaries and structure of the moral community. This is because, by focusing on an agent's evaluative stance, answerability is meant to channel third parties' attitudes, such as resentment, admiration and gratitude that are made appropriate by the expression of the agent's stance. Following Peter Strawson (1962), these reactive or participatory attitudes are treated as constitutive for having a sense of oneself and others as (full) members of the same moral community (cf. Schoemaker and Tognazzini 2014). The thought is that by giving voice to our disappointment or delight with other people's actions and attitudes, judgements about responsibility embedded in reactive

attitudes enable us to articulate a notion of moral agents participating in a shared practice or network of practices, within which it is sufficiently clear, in principle, what might count as offensive as opposed to respectful or friendly. Here, knowing what one is doing has a decidedly normative component. It involves the assumption of a degree of normative competence that would make a reaction in terms of resentment or gratitude both deserved and intelligible to its intended target.

The kind of internalism about reasons for action endorsed by proponents of the revised knowledge condition for responsibility is at odds with the central communicative and interactive function that responsibility assessments in the form of reactive attitudes are expected to fulfil. Pursued to its conclusion, the unbridgeable gap between internal and external reasons leads to two possible options. The first is to limit the scope of appropriate resentment to what is condemnable by the agent's own lights, which opens the door to a generalised scepticism about moral responsibility. The second is to abandon concerns about fairness in allocating moral responsibility as misplaced. Let us briefly consider each of these options in turn.

The first option flows from confining the reasons for action available to an agent to items within her existing motivational set and developments that can be plausibly derived from it. The thought that we are not entitled to ask others to do things that would make no sense from their own perspective yields intuitive support to constrain third-personal responsibility assessments in this way. Furthermore, since the revised knowledge condition should not include elements of reflection or discretion, fairness to an agent as a possible target of a negative reactive attitude, such as resentment would require that she actually knows better. And so, knowing what one is doing warrants (negative) responsibility to the extent that it involves knowing that what one is doing is wrong, impermissible, offensive etc. As Gideon Rosen points out in a related discussion (2004, p. 307):

One is responsible for the act done from ignorance only if one is independently responsible for something else...this entails that *the only possible locus of original responsibility is an akratic act...* Our first sin must be a knowing sin – a sin done in full knowledge of every pertinent fact or principle.

From this observation, Rosen argues in favour of a generalised scepticism about moral responsibility in the sense of warranted doubt about resentment, a central reactive attitude, ever being apt. Once the practice of holding each other responsible is rooted into the

respective motivational sets of individual agents that may or may not converge toward a shared set of normative expectations, it becomes difficult to see how such a sceptical conclusion could be resisted, on grounds of interpretative fairness.

The second option available to proponents of the revised knowledge condition is to diffuse concerns about fairness as remnants of the rejected control condition. This approach is taken up by Hieronymi who argues that ‘criticising resentment because it is unfair burden on the one resented is like criticising a belief by pointing out that it has bad consequences’ (2014, p.30). According to Hieronymi, the question of fairness does not arise with respect to distrust either. This is because distrust tracks a person’s present unreliability, independently of whether this person has neglected the opportunity of becoming reliable in the past or could be persuaded to become more reliable in the future. Following this line of reasoning, it is concluded that moral demands are not pedagogical or ‘custom fit to the recipient’, ‘they stand unyielding in the face of an inability to meet them; [...and] the reactions we have when we regard one another as responsible are neither punitive nor voluntary; [...] the reactions need not include a commitment to the claim that their target could have avoided wrongdoing by trying harder’ (p.35).

To recap, this paper explored the relationship between two related trends in rethinking responsibility, downplaying the control condition in terms of voluntariness and choice while emphasising the knowledge condition in terms of first-order evaluative judgment and belief. Looking at the communicative and interactive functions that responsibility assessments are still meant to fulfil on conceptions that exemplify these two trends, there are reasons to doubt the prospects of answerability without discretion or reflection.

## References

Anscombe, G.E.M. (1963). *Intention*, 2nd ed. Oxford: Blackwell.

Arpaly, N. (2003). *Unprincipled Virtue: An Inquiry into Moral Agency*. Oxford: Oxford University Press.

Deery, O., Bedke, M. and S. Nichols (2013). Phenomenal Abilities: Incompatibilism and the Experience of Agency. In Schoemaker, D. (ed.), *Oxford Studies in Agency and Responsibility*, vol. 1; pp. 126–151.

- Fisher, J.M. and M. Ravizza (1998). *Responsibility and Control. An Essay on Moral Responsibility*. Cambridge: Cambridge University Press.
- Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy* 66: 829–839.
- Hieronymi, P. (2014) Reflection and responsibility. *Philosophy & Public Affairs* 42: 3–41.
- Levy, N. (2005). The good, the bad, and the blameworthy. *Journal of Ethics and Social Philosophy* 1: 1–16.
- Levy, N. (2014a). Consciousness, implicit attitudes and moral responsibility. *Noûs* 48: 21–40.
- Levy, N. (2014b). Neither fish nor fowl: implicit attitudes as patchy endorsements. *Noûs* 00: 1–24. doi: 10.1111/nous.12074.
- Radoilska, L. (2013). *Addiction and Weakness of Will*. Oxford: Oxford University Press.
- Rosen, G. (2004). Scepticism about moral responsibility. *Philosophical Perspectives* 18: 295–313.
- Scanlon, T. (2008). *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, MA: Belknap Press.
- Schoemaker, D. and Tognazzini, N. (eds.) (2014) *'Freedom and Resentment' at 50: Oxford Studies in Agency and Responsibility*, vol. 2.
- Smith, A. (2005). Responsibility for attitudes: activity and passivity in mental life. *Ethics* 115: 236–271.
- Smith, A. (2008) Control, responsibility, and moral assessment. *Philosophical Studies* 138: 367–392.
- Strawson, P. (1962). Freedom and resentment. *Proceedings of the British Academy* 48:1–25.
- Talbert, M. (2013). Unwitting wrongdoers and the role of moral disagreement in blame. In Schroeder, M. (ed.), *Oxford Studies in Agency and Responsibility*, vol. 1; pp. 225–245.
- Wallace, R.J. (1994). *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.

Wallace, R.J. (2006). *Normativity and the Will*. Oxford: Oxford University Press.

Widerker, D. and McKenna, M. (eds.) (2003). *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*. Aldershot: Ashgate.

Williams, B. (1981). Internal and external reasons. In B. Williams. *Moral Luck*. Cambridge: Cambridge University Press; 101–113.