Author-Approved Version. October 2020. Forthcoming in: Luca Ferrero (ed.) *The Routledge Handbook of the Philosophy of Agency*

PATHOLOGIES OF AGENCY

Lubomira Radoilska
Department of Philosophy, University of Kent
Canterbury, Kent, CT2 7NF, UK
L.V.Radoilska@kent.ac.uk

ABSTRACT: This chapter aims to distinguish between pathologies of agency in the strict sense and mere sources of impediments or distortion. Expanding on a recent notion of necessarily less-than-successful agency, it complements a mainstream approach to mental disorders and anomalous psychological conditions in the philosophy of mind and action. According this approach, the interest of such clinical case studies is heuristic, to differentiate between facets of agency that are functionally and conceptually separate even though they typically come together. Yet, in the absence of independent criterion for a pathology of as opposed to inner obstacle to agency, this heuristic is at risk of becoming circular or uninformative, falling back on a clinical diagnosis it is meant to take as a starting point only. The chapter develops such a criterion and shows how it could work tracking agential achievement across two core dimensions of agency: planning and responsiveness to reasons. The discussion concludes with some implications on assessing decisional capacity and safeguarding agent autonomy in psychiatric settings.

KEYWORDS: achievement; agency; decisional capacity; mental disorders; necessarily less-than-successful agency; pathologies of agency; planning; psychiatry; responsiveness to reasons

Mental disorders and anomalous psychological conditions play a crucial role in defining the nature and scope of agency. The standard approach describes a particular set of symptoms as a point of contrast where some core features identified in paradigm cases are present while others are clearly absent. For instance, the literature on psychopathy (e.g., Schramme 2014) often points to the

combination of single-minded pursuit of personal goals with lack of empathy leading to casual disregard for the interests of others. On this picture, psychopathy instantiates a distinctive shape of agency where some reasons can be as integrated as in central cases of successful planning whereas others get no purchase. Alien Hand Syndrome (unbuttoning own shirt with one hand is vigorously fought off with the other, Pacherie 2007) is a further example. It shows how intentional agency might unfold when the ability for basic bodily movements becomes fragmented. Finally, people with depression may be motivated to act in ways they manifestly disvalue (Radoilska 2013a). The experience is both alienating and disturbing. It decouples first-personal knowledge of the reasons for which one acts from the self-understanding that acting on one's own reasons is supposed to secure.

These examples show that close attention to the phenomenology of specific disorders and anomalous conditions helps differentiate between agential abilities that textbook cases of action bundle together. This is the first step of a heuristic that explores the possible interactions between facets of agency that typically come together but are functionally and conceptually separate. The payoff is twofold. First, we gain a clearer insight into manifestations of agency beyond straightforward intentional actions, including negative, omissive, and second-order exercises, such as permitting, preventing, facilitating, or contravening (Alvarez 2013). Second, we can better understand and assess competing attributions of responsibility by mapping them onto different patterns of agency (Shoemaker 2015).

At the same time, this heuristic might become circular and uninformative. The danger comes from sticking to the phenomenology so closely that no distinction is made between pathologies of agency and mere sources of impediments and distortion. Without such distinction, the lessons learnt about diverse facets and shapes of agency in the context of mental disorders could become unreliable. For contours of agency might be unhelpfully merged with its circumstance. To forestall this danger, we need to complement the mainstream heuristic. This chapter will try to do so by

expanding on the notion of necessarily-less-than successful agency (Radoilska 2013b). A major advantage of this proposal is to do justice of 'achievement' as unifying idea that underpins all credible manifestations of agency (Bradford 2015). By focussing on how and why achievement might be affected by different mental disorders and anomalous psychological conditions, it becomes possible to avoid the circularity invited by mainstream heuristics: falling back on a clinical diagnosis to pinpoint a pathology *of*, as opposed to an obstacle *to*, agency.

To see this, the notion of necessarily-less-than successful agency will be employed to identify and explore specific pathologies within two dimensions of agential achievement: intention and responsiveness to reasons.

Intention, Practical Knowledge and Planning

Intention and intentional action are central manifestations of agency, to the point of obfuscating alternatives. To appreciate the attractiveness of thinking about agency through the lens of intentional action, let us consider the influential account of Anscombe (1963). On this account, intentional actions are the subject of a distinctive kind of direct first-personal knowledge—unmediated by observation, inference, or reflection—, which should not be equated with introspection or general awareness of one's own mental states. For such a move would reduce the agent's perspective to that of a well-placed, 'inside' observer and sever the 'mind-world' relation that comes with intention as stretching out into the world (Moran and Stone 2009). Yet both perspective and relation are irreducible aspects of the relevant kind of fundamentally practical first-personal knowledge. The practical nature of knowing one's intentional actions has direct implications regarding success in action. The success criteria are set out by the agent's account of what they are doing. When this account differs from what actually happens, the 'words impugn the facts' as Anscombe puts it, not the other way around. For the failure to conform betrays an error in performance, not an error in judgment. Importantly, while such errors in performance are

rare, agents often miss on their overall objectives. This observation shows that two kinds of practical knowledge may come apart: knowing what one is doing intentionally at present, on the one hand, and generalised know-how that enables agents to plan and coordinate beyond the performance of discrete intentional actions, on the other (but see Setiya 2008 on the intimate relation between these two kinds of knowledge). Satisfying the criteria of success set out with respect to the first kind of practical knowledge might not suffice overall. On these criteria, individual actions can be successful independently of whether they add up to a defensible strategy that brings them together. As a result, we cannot tell between a savvy agent who keeps an eye on longer-term objectives and one who gets bogged down in unnecessary detail. If agential achievement is to be assessed in a meaningful way, another set of criteria, linked to the second kind of practical knowledge—generalised know-how exercised over time, across different situations—becomes indispensable.

This insight has been taken forward by the planning theory of intention (Bratman 1987; 2007). According to this view, an intention has an irreducibly dual function: it guides the performance of individual actions and it connects with other intentions to enable the pursuit of complex, temporarily extended projects. Looking at intentional agency from this integrative perspective, we can gauge a new set of success criteria to complement the idea of practical knowledge. Roughly, success is a matter of striking a balance between the requirements of coherence across plans and consistency over time with the need for plans to be flexible in response to changing circumstances. Let us call this task 'practical rationality' to distinguish it from 'practical knowledge.'

With this distinction in mind, let us focus on a puzzling pathology of agency where the notion of necessarily-less-than-successful agency comes into its own.

Jill has an important presentation to make next day early in the morning. Before heading back home after work, her colleagues invite her to join them for a drink. She decides to go out but have no more than two drinks as she knows that drinking any more will affect the quality of her

performance next morning. As the evening progresses, she is offered a third drink, which, after a brief moment of hesitation, she takes. Against her better judgment, she ends up having a fourth drink as well.

This kind of scenario abounds in the literature on weakness of will. There are two mainstream accounts of what goes wrong with Jill. According to Mele (1987), the problem is Jill's acting against her better judgment at the time of action. She exhibits weakness of will on this occasion, which constitutes a failure of practical rationality. We don't need to know anything further about her as an agent to reach this conclusion. All relevant information is already contained in this description.

According to Holton (2009), instead, the scenario hints at a possible case of weakness of will if Jill's decision to have only two drinks is a personal policy adopted in light of previous trouble with alcohol. To be deemed as weak-willed and therefore practically irrational, Jill's failure to see through her good intentions must be representative of a pattern rather than a one-off.

The two mainstream accounts derive from the two perspectives on what counts as success in action. While the first focuses exclusively on the standard set out by a specific intention, the second also considers the diachronic implications of any particular failure. These differences notwithstanding, both accounts concur on weakness of will being a failure of agency in contrast to compulsion, a pathology of agency whose distinguishing feature is loss of agential control. The thought is that a pathology of agency would fall outside the scope of criticisable irrationality, which makes failures of agency, such as weakness of will intelligible (Davidson 2001). In other words, pathological agents do not stand a chance to succeed. And so, their irrationality is regrettable rather than criticisable.

Although initially appealing, the demarcation line between failing and pathological agency that compulsion provides is ultimately misleading. For it assumes a one-dimensional notion of success in action as bringing about a desirable result—be it at a snapshot or in a temporarily extended frame. Yet on closer inspection, weakness of will and compulsion have a central common feature

that helps elucidate and differentiate among these closely related phenomena: being necessarily less-than-successful with respect to a particular strand of one's goal-directed conduct rather than one's overall agency. This feature boils down to neither suboptimal behaviour, nor a straight failure. Instead, it points to a distinctive structure of agency, the manifestations of which cannot be fully successful to the extent that they arise at all (Radoilska 2013b).

Unlike the two mainstream accounts, the notion of necessarily less-than-successful agency implies a more complex model of action as actualisation, where success is defined across two complementary dimensions: production (bringing about an effect) but also assertion (an agent's articulating a particular commitment of theirs). Only when these two dimensions are well-aligned is an action successful on its own terms. By contrast, when they are misaligned in a distinctive and sustained way, rather than just coming apart, the ensuing actions are necessarily less-than-successful. Returning to compulsion and weakness of will, the misalignment that transpires in both takes the following form: each is successful as production to the extent that it is unsuccessful as assertion. To illustrate with the scenario we considered, what Jill does is successful as production (she gets to have more than two drinks) to the extent that it is unsuccessful as assertion (it goes against her better judgement and/or policy). There is an underlying conflict between valuing and intending with respect to drink that we may call Jill's original akratic moment. If this conflict is not addressed but instead keeps coming back, it eventually solidifies into a necessarily less-than-successful strand of Jill's agency.

Is this a case of weakness of will or compulsion? Does it amount to a pathology as opposed to failure of agency?

Addressing the latter issue first, we can see that partial, yet unavoidable failure is constitutive of pathological agency. Far from placing agents outside the space of reasons, where the charge of being practically irrational is no longer apt, pathologies of agency provide this charge with a specific focus. This explains why compulsion and the binary notion of control vs. irresistibility it implies

cannot give us a reliable cutting point. In fact, many cases, such as the preceding scenario would satisfy the criteria of necessarily less-than-successful agency that underpin both phenomena. So, the differences we may find will be of degree rather than kind. When it is more illuminating to conceive the pathology manifested in terms of addiction rather than weakness of will, the relevant actions are experienced as particularly frustrating attempts to resolve the recurring conflict between valuing and intending that effectively extend the original akratic moment to the detriment of other, potentially successful pursuits of one's agency. This feature is at the heart of debates about the extent to which addictive behaviours can be treated as responsive to reasons (Poland and Graham 2011). In the next section, we will explore how and why responsiveness to reasons might be affected by various pathologies of agency, leading to differential appraisals of agents as members of the moral community.

Responsiveness to Reasons, Answerability and the Moral Community

The Anscombian notion of practical knowledge sets out an immediate test of whether an action is successful. Knowing what one is doing without having to observe oneself or infer from prior experiences provides the description under which what one is doing is intentional. For instance, I don't need to look up at what I am doing to know that I am currently opening a window, if that is an intentional action of mine rather than something, I find myself doing. Nor do I need to work out what I am up to in similar cases. Thus, an agent's performance is assessed, in the first instance at least, against such a description. What makes it special is that it designates an action as undertaken in the light of the agent's reasons as opposed to behaviours understood in causal terms. Knowing what one is doing in the relevant practical sense rests on knowing why one is doing it. My giving account of what I am doing is unlike my reporting on what is happening around me—with or without my causal involvement. This is the guiding idea behind the Anscombian question 'Why?'.

It is helpful to contrast the responses this question is meant to elicit with action explanations afforded by the so-called 'reasons why' (Dancy 2000), where the language of reasons is used without referring to any considerations, in the light of which an agent acts. Examples include my failing to turn up to a party because I forgot about it or because I am too shy. Neither figures in my reasoning whether to go the party and, yet, either can rationalise my not being there. In this respect, they both could present bona fide 'reasons why' for my staying away from the party. In so doing, however, they would mark out what I did as behaviour, where the Anscombian question 'Why?' is denied application. 'Reasons why' do not warrant the kind of knowledge agents have of actions they perform in the light of reasons. What is lost is the immediacy and certainty of practical knowledge: knowing what one is doing, as Hornsby puts it, 'without recourse to further knowledge' or 'just like that' (2013, p.16).

The expectation of an immediate, unbreakable bond between actions done for a reason and actions of which agents have practical knowledge is put to the test by apathy or auto-activation deficit syndrome (AAD), a neurological disorder where self-generated voluntary and purposive actions are virtually absent, while externally driven behaviours are normally executed. As Levy (2012, p. 590) explains:

> This syndrome consists in a loss of spontaneous activation in three different domains: behaviour, cognition and emotion. Patients tend to remain quietly in the same place or position all day long, without speaking or taking any spontaneous initiative. When questioned, patients express the feeling that their mind is empty when they are not stimulated… [however] when solicited, patients can produce relevant answers and behaviours.

The following case studies illustrate the range of behaviours where apathy is at work:

- Patient A spent 45 minutes with his hands on a lawn mower, totally unable to initiate the act of mowing. The block disappeared instantaneously when his son told him to move (Laplane and Dubois 2001).

- Patient B developed a hobby of collecting broken TVs. His collection filled up the family flat, spilt over into the common premises of the building risking eviction. B showed awareness of the consequences of his new hobby, denied interest in his collection and could not explain why he nevertheless kept on bringing more and more items (Levy 2015).
- Research subject C saw the keys for the researcher's car on his desk, took the keys, got into the car, did a few rounds and returns to the room. When asked what she did, C was able to describe accurately her actions but could not provide any reason for them other than seeing the researcher's car keys on the desk (*ibid*).

What makes these cases unsettling is the contrast between the lucidity with which people with apathy are able to report on what they are doing without having any inkling as to the reasons motivating their actions. They seem to be responding to cues or incentives in the environment as opposed to reasons they can recognise as their own. The difference between these two kinds of responses has significant implications for attributions of responsibility and moral appraisal.

To appreciate this, consider the central place of notions, such as answerability, in mainstream conceptions of responsibility. In Hieronymi (2014), the practice of responsibility is essentially about the way in which an agent settles questions like whether to undertake a particular course of action, maintain an attitude or a relationship, or revisit an existing commitment. In all relevant cases, the agent is answerable to the Anscombian question 'Why?'. This question tracks the reasons in the light of which the agent acted or refrained from acting, maintained or revisited attitudes, commitments, or relationships of hers.

Answerability provides an attractive model able to account for the interpersonal significance of allocating responsibility as opposed to a disinterested study of why people do the things they do. For it speaks directly to the idea of a counterfactual conversation with those affected by one's actions or an internal dialogue that mainstream conceptions of responsibility build upon (e.g. McKenna 2012, Wallace 1994). In such a communicative setting, giving reasons for one's actions

plays a key role in being and holding responsible (Smith 2007). It involves anticipating challenges, but also showing due concern for others' interests and perspectives. In this way, reason-giving demonstrates an agent's secure belonging to the moral community where the so-called reactive attitudes, such as resentment or gratitude, the backbone of being and holding responsible, can be appropriately addressed (Strawson 1962).

As seen earlier in cases of apathy, competently executed behaviours may not be responsive to reasons but only sensitive to incentives. Arguably, such behaviours would not qualify as appropriate ground for reactive attitudes, nor would reflect negatively on the moral standing of apathic agents: Patient B's family did not reproach him for risking eviction, nor asked him to abandon his collection of broken TVs. Instead, they sought advice from an AAD specialist to see how best to manage the situation.

Although in other cases the divide between reasons and incentives might not be as clear-cut, the same logic would apply to separating out responsible conduct from conduct that should be exempt from responsibility. For instance, according to an influential account of addiction as incentive-sensitisation (Berridge and Robinson 2011), blame is unsuitable response to addictive behaviours since they are not reasons-responsive. Instead, social effort should be directed at restructuring everyday environments so that people with addictions are not readily exposed to the incentives they have been sensitised to. In this respect, addiction is seen as an obstacle to personal agency rather than a pathology of it: there are situations where a person with a particular addiction cannot be expected to fare well. So, we better make it easy for her to avoid such situations rather than call on her to demonstrate uncommon strength of will and then berate her for not doing so.

Importantly, when mental disorders and anomalous psychological conditions are conceived as obstacles to responsiveness to reasons within a particular sphere or context, they do not impact on the perception of the agents affected as members of the moral community. Research on the moral commitments and agency of people with autism (Kennett 2002) and intellectual disabilities

(Shoemaker 2009) is a case in point. While in both instances, access to morally relevant reasons can be seen as impeded, there remain alternative pathways to reasons for caring for others and respecting their interests. This enables full participation in the practice of being and holding responsible in stark contrast with agent appraisal in cases of psychopathy. As Kennett (2002) and Shoemaker (2009) highlight, there is a qualitative difference in the way morally relevant reasons are present in practical reasoning in these contrasting cases. For instance, a person with a mild to moderate intellectual disability might be initially unable to see a particular course of action as morally required. However, once the rationale for it is brought to her attention, she would appreciate its binding force. By contrast, psychopaths are effectively able to contemplate morally relevant reasons. In this sense, they have unimpeded access to what morality requires of them. Yet, they do not treat these requirements as binding. Their response to moral reasons is warped rather than obstructed. As evidenced by the popular perception of psychopaths as amoral, or permanent outsiders to the moral community, indifference to the action-guiding aspect of morality grounds a particularly robust negative moral appraisal. Thus, psychopaths are treated as appropriate target for blame and resentment even though, ex hypothesi, they are not expected to engage in the kind of meaningful, reciprocal exchange that instantiates responsiveness to reasons proper. In this respect, psychopathic indifference to reasons for action that are nevertheless successfully identified bears the hallmarks of necessarily less-than-successful agency we discussed in the previous section. As in akratic action, the response to other-regarding reasons afforded by psychopathy cannot be fully successful to the extent that it takes place at all: for misrecognition is the mode of their cognition. The pathology of agency here does not consist in having outside one's practical competence tasks, whose performance is typically taken for granted by others, as was the case with incentive sensitisation. Instead, the crux of the matter is that the stable, self-fulfilling mechanism underpinning necessarily less-than-successful agency makes alternative approaches to agential achievement, to which success and failure are equally open, riskier and less attractive as a result.

Recent work on rationalisation shows how responsiveness to reasons in non-clinical populations might also exhibit a relevantly similar mechanism maintaining necessarily less-than-successful projects over time. As Schwitzgebel and Ellis (2017, p. 170) point out:

> Rationalisation occurs when a person favours a particular conclusion as a result of some factor (such as self-interest) that is of little justificatory epistemic relevance. The thinker then seeks an adequate justification for that conclusion but the very factor responsible for her favouring it now biases how the research for justification unfolds. As a result of an epistemically illegitimate investigation, the person identifies and endorses a justification that makes no mention of the distorting factor that has helped guide her search.

There is a dissonance here between three categories of reasons: actual motivating reasons, 'reasons why' which explain the process of rationalisation, and reasons which the agent ultimately professes as those in the light of which she made up her mind. Rationalisation is especially pernicious as it echoes the self-fulfilling mechanism we observed in other necessarily less-than-successful manifestations of agency. It effectively papers over the dissonance between different categories of reasons, making its detection and eventual resolution extremely unlikely. The upshot has direct implications for the practice of being and holding responsible. As Schwitzgebel and Ellis (2017, p. 171) observe, rationalisation not only 'obstructs the critical evaluation of one's own reasoning'; in addition, 'it impedes the productive exchange of reasons and ideas among well-meaning interlocutors'.

Implications for Decisional Capacity and Autonomy in the Context of Mental Disorder

The preceding discussion showed that neither mental disorders nor abnormal psychological conditions necessarily imply a pathology of agency. Sometimes, their impact on agency is best understood in terms of obstacle to sidestep or constraint to overcome via targeted strategies. This conclusion finds further support in the literature on decisional capacity and personal autonomy in

psychiatric contexts. For instance, many authors draw attention to the harmful nature of implicit assumptions about rationality, according to which a diagnosed psychiatric condition is enough to put a question mark on a person's decisional capacity (Culvert and Gert 2004; Bolton and Banner 2012; Bortolotti 2013). In response to this, international psychiatric practice is developing toward robust and clear policies embedding the first-personal perspective and experiences of service-users at the heart of treatment (Widdershoven and Abma 2012; Potter 2013). The ambition is to prevent failures to recognise as reasons-responsive projects and commitments that are of great personal significance merely because they might seem unusual or unappealing from the perspective of a clinician. Yet, in the absence of a positive notion of what constitutes a pathology of agency, this mainstream approach can backfire in clinical settings. As Jamison (1995) and Radden (2012) poignantly illustrate in the context of bipolar disorder, it is equally important to address threats to personal autonomy that derive from having one's behaviours that are not reasons-responsive treated as though they were. The unsympathetic imposition of penalties for such behaviours is an immediate issue. Examples include crippling loans undertaken during a manic episode, to fund out-of-character hobbies. A deeper concern, however, is that misallocating responsibility in this way would generalise any pathological strands of agency already present (Radoilska 2015). The concept of a necessarily less-than-successful structure could be of help here. For it allows us to pinpoint, clinical diagnosis notwithstanding, the kind of ambivalent, ultimately self-defeating achievements that may not be protected in the name of agent autonomy.

Further reading

- Fulford, K.W.M., Davies, M., Gipps, R.G.T., Graham, G., Sandler, J.Z., Stranghellini, G. and T. Thornton (eds.) 2013. *The Oxford Handbook of Philosophy and Psychiatry*. Oxford: Oxford University Press.

The essays included in this collection offer a comprehensive overview of how and why mental disorder might impact on different aspects of agency. It brings together conceptual analysis,

empirical research and clinical studies to shed light on philosophical and normative issues arising from psychiatric practice.

- Radoilska, L. (ed.) 2012. *Autonomy and Mental Disorder*. Oxford: Oxford University Press.

This anthology is the first book-length multidisciplinary inquiry into the nature and scope of personal autonomy in the context of mental disorder. It helps challenge tacit assumptions in the philosophy literature, according to which mental disorders are primarily threats to autonomous agency.

- Radoilska, L. 2013. *Addiction and Weakness of Will*. Oxford: Oxford University Press.

This monograph introduces the notion of necessarily less-than-successful action to qualify the kind of criticisable irrationality that underpins both addiction and weakness of will. It argues that being necessarily less-than-successful rather than a symptom or clinical diagnosis is what distinguishes pathologies from inner obstacles to agency.

- Schramme, T. (ed.) 2014. *Being Amoral: Psychopathy and Moral Incapacity*. MIT Press.

The collection addresses the question of what makes the moral agency of psychopaths wanting. Against the grain, the contributors show that psychopathy is best understood as a cluster of deficits affecting different moral capacities rather than a uniform lack of empathy toward others.

- Shoemaker, D. 2015. *Responsibility from the Margins*. Oxford: Oxford University Press.

The monograph develops a pluralist conception of responsible agency which can be broken down into three main facets: answerability, accountability and attributability. Different mental disorders and abnormal psychological conditions are shown to preclude some facets of responsible agency while at the same time leaving others intact.

Biographical note

Lubomira Radoilska is a Senior Lecturer in Philosophy, University of Kent, UK. She works on issues at the intersection of philosophy of action, ethics and epistemology. Radoilska is the author of *Addiction and Weakness of Will* (OUP, 2013) and editor of *Autonomy and Mental Disorder* (OUP, 2012).

References

Alvarez, M. 2013. Agency and Two-Way Powers. *Proceedings of the Aristotelian Society* 113(1): 101–121.

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

Berridge, K. and Robinson, T. 2011. Drug Addiction and Incentive Sensitization. In Poland, J. and G. Graham (eds.). *Addiction and Responsibility*. Cambridge, MA: MIT Press; 21–53.

Bolton, D. and Banner, N. 2012. Does Mental Disorder Involve Loss of Personal Autonomy? In Radoilska, L. (ed.) *Autonomy and Mental Disorder*. Oxford: Oxford University Press; 77–99.

Bortolotti, L. 2013. Rationality and Sanity: The Role of Rationality Judgments in Understanding Psychiatric Disorders. In K.W.M. Fulford, M. Davies et al. (eds.) *The Oxford Handbook of Philosophy and Psychiatry*. Oxford: Oxford University Press, 480–496.

Bradford, G. 2015. *Achievement*. Oxford: Oxford University Press.

Bratman, M. 1987. *Intention, Plans and Practical Reason*. Cambridge, MA: Harvard University Press.

Bratman, M. 2007. *Structures of Agency*. Oxford: Oxford University Press.

Culvert, C.M. and Gert, B. 2004. Competence. In J. Radden (ed.) *The Philosophy of Psychiatry*: A Companion. Oxford: Oxford University Press; 258–270.

Dancy, J. 2000. *Practical Reality*. Clarendon Press.

Davidson, D. 2001. How Is Weakness of the Will Possible? In D. Davidson, *Essays on Actions and Events*. Oxford: Clarendon Press; 21–42.

Hieronymi, P. 2014. Reflection and Responsibility. *Philosophy & Public Affairs* 42: 3–41.

Holton, R. 2009. *Willing, Wanting, Waiting*. Oxford: Oxford University Press.

Hornsby, J. 2013. Basic Activity. *Proceedings of the Aristotelian Society*, Suppl. Vol. 87:1-18.

Kennett, J. 2002. Autism, empathy and moral agency. *Philosophical Quarterly* 52 (208): 340-357.

Jamison, K. R. 1995. An Unquiet Mind. New York: Vintage Books Random House.

Laplane, D. and Dubois, B. 2001. Auto-Activation Deficit: A Basal Ganglia Related Syndrome. *Movement Disorders* 16: 810–814.

Levy, R. 2012. Apathy: A Pathology of Goal-Directed Behaviour. A New Concept of the Clinic and Pathophysiology of Apathy. *Revue Neurologique* 168: 585–597.

Levy, R. 2015. Apathy: Anonymised Case-Studies. *Neuroscience and Philosophy*. Paris-Sorbonne Universités, 27 Nov.

McKenna, M. 2012. *Conversation and Responsibility*. Oxford: Oxford University Press.

Mele, A.R. 1987. *Irrationality: An Essay on Akrasia, Self-Deception and Self-Control*. New York: Oxford University Press.

Moran, R. and Stone, M. 2009. Anscombe on the Expression of Intention. In Sandis, C. (ed.) *New Essays on the Explanation of Action*. Palgrave Macmillan; 132–168.

Pacherie, E. 2007. The Anarchic Hand Syndrome and Utilization Behaviour: A Window onto Agentive Self-Awareness. *Functional Neurology* 22 (4):211–217.

Poland, J. and Graham, G. (eds.) 2011. *Addiction and Responsibility*. Cambridge, MA: MIT Press.

Potter, N. N. 2013. Empathic Foundations of Clinical Knowledge. In K.W.M. Fulford, M. Davies et al. (eds.) *The Oxford Handbook of Philosophy and Psychiatry*. Oxford: Oxford University Press, 293–306.

Radden, J. 2012. Privacy and Patient Autonomy in Mental Healthcare. In Radoilska, L. (ed.) *Autonomy and Mental Disorder*. Oxford: Oxford University Press; 123–142.

Radoilska, L. 2013a. Depression, Decisional Capacity, and Personal Autonomy. In K.W.M. Fulford, M. Davies et al. (eds.) *The Oxford Handbook of Philosophy and Psychiatry*. Oxford: Oxford University Press, 1155–1170.

Radoilska, L. 2013b. *Addiction and Weakness of Will*. Oxford: Oxford University Press.

Radoilska, L. 2015. Autonomy in Psychiatric Ethics. In Sadler, J., van Staden, G. W. and Fulford, K. W. M. (eds.) *The Oxford Handbook of Psychiatric Ethics*. Oxford: Oxford University Press; 354-371.

Schramme, T. (ed.) 2014. *Being Amoral: Psychopathy and Moral Incapacity*. MIT Press.

Schwitzgebel, E. and Ellis, J. 2017. Rationalization in Moral and Philosophical Thought. In: Bonnefon, J.F. and B. Trémolière (eds.). Moral Inferences. Routledge, New York: 170–190.

Setiya, K. 2008. Practical Knowledge. *Ethics* 118 (3): 388–409.

Shoemaker, D. 2009.Responsibility and Disability. *Metaphilosophy* 40(3/4): 438-461.

Shoemaker, D. 2015. *Responsibility from the Margins*. Oxford: Oxford University Press.

Smith, A. 2007. On Being Responsible and Holding Responsible. *The Journal of Ethics* 11: 465-484.

Strawson, P. 1962. Freedom and Resentment. *Proceedings of the British Academy* 48: 1–25.

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

Widdershoven, G.A.M. and Abma, T.A. 2012. Autonomy, Dialogue and Practical Rationality. In Radoilska, L. (ed.) *Autonomy and Mental Disorder*. Oxford: Oxford University Press; 217–232.