

Un-willed Beliefs:
An Essay on Voluntariness and Doxastic
Voluntarism

by
Benjamin Davis

Submitted for the degree of MPhil Stud in Philosophy
UCL
August 2022

I, Benjamin Davis, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

אמר להו רבי יוסי תורה מבחוץ ואנו מבפנים
אמר להן רבן [שמעון בן] גמליאל ניעיילינהו

Rabbi Yossi said to his colleagues, “the Torah is outside whilst we’re inside the academy!”
Rabban [Shimon ben] Gamliel said to them, “let’s bring it in.”

Babylonian Talmud, *Horayot*, 13b

The world is independent of my will.

Wittgenstein, *Tractatus Logico-Philosophicus*, 6.373

Abstract

Doxastic involuntarism is somewhat of an epistemological orthodoxy. Contra this orthodoxy, I argue that doxastic voluntarism is *routinely* true. The aim is to argue for this by exploring the nature of the voluntary. Following Hyman, I argue that our apprehension of voluntariness has been subject to serious error. In particular, our notion of what is voluntary has become conflated with notions about what it is to act as such. This is because theories of action confounded *acts of will* with voluntariness. This diminishes us philosophically because in consequence, we understand neither action *simpliciter* nor voluntariness on their own terms. I defend and argue for Hyman's position that voluntariness is a concept whose primary dominion spans the *ethical* dimension of human conduct, whereas the question of what it is to act concerns *agency* as such; what it is for a power to produce or cause change, and how action differs from mere movement. By restoring voluntariness back to the ethical axis of human conduct, we achieve a clearer view of its nature. Voluntariness is destroyed by ignorance and compulsion. Moreover, only the doings and undergoings of *doxastic* and *evaluative* creatures are properly subject to predications of voluntariness and involuntariness.

I use this to articulate a new account of doxastic voluntarism, because epistemology too has been infected by this error. Voluntarism is true just in case the doxastic ignorance and compulsion conditions *don't* apply. Upon elucidating these conditions, I argue that generally, these conditions don't apply when we form and hold beliefs. This is true whether we are doxastic agents or patients. By abjuring notions of will from doxastic voluntarism and involuntarism, that is, by *un-willing beliefs*, we gain a clear picture of the relation between voluntariness and belief. The picture informs us that doxastic voluntarism is true, and routinely so.

Impact Statement

This thesis argues for and defends a new way to understand doxastic voluntarism and doxastic involuntarism. The proposal itself is philosophically significant (even if my arguments are ultimately unsuccessful) because it reveals a clearer way of grasping the relation between voluntariness and belief. If my contentions are correct, the impact should be felt: my study will contribute a novel way to overturn a philosophical orthodoxy. That is, it will show not only that doxastic voluntarism is conceptually possible, but that our beliefs are *ordinarily* acquired and held by voluntary means. Based off this account of doxastic voluntarism, I also argue for a corresponding account of doxastic responsibility for in/voluntary beliefs. Therefore, my thesis will benefit philosophers working on epistemology, responsibility, and the philosophy of mind and action. It will also impact other areas of philosophy where the concept of belief is central such as the philosophy of religion and the philosophy of psychiatry.

My research can also have implications for philosophically contiguous areas of concern, which are nevertheless outside academia. For example, I am excited to see how my findings will develop into better understanding some psychiatric conditions which are characterised by a set of beliefs a subject possesses. It will also speak to therapeutic practices which stress the importance to the subject of ‘taking responsibility’ for their beliefs. But this, of course, depends on future work and research though the foundations are here in this thesis.

Acknowledgements

The past so many years, I have been fortunate to have had my major philosophic training in the philosophy department at UCL. It has been formative. Any philosophical progress or development I can track within me, if there be any at all, I know is owed largely to my time in the department. I have been taught by a number of excellent teachers and I have interacted with excellent postgraduate students. (The last of which was so missed during those lockdown periods). I want to collectively thank the postgraduate cohort for their (our), I think, enviable collegiate and collaborative spirit. In the same vein of collective thanks, I shan't name them all here, but there are faculty members who have along the way, been integral to my education. I thank them all.

I must make explicit mention of two of my teachers, my supervisors, Prof. John Hyman and latterly, Prof. Lucy O'Brien. I largely credit Lucy with instigating a transformation in my self-understanding with regards to what I'm doing as a philosophy student. (That already is a sentence replete with material apt for Lucy's analysis: it includes an attitudinal change, self-conscious awareness of that attitudinal change, and a 'doing'). If I were to mark a precise origin in this change, it would be taking Lucy's Advanced Class in the Philosophy of Mind (2021). I began to understand that my interest in doxastic attitudes and ethics of belief need not be an interest confined to epistemology, ethics, and their overlap. Actually, to get to the depths I wanted to, I realised that this interest can (must?) incorporate an exploration into mind and action. In particular, the nature of doxastic agency and mental acts. Thus, even if it isn't obvious to the reader of this thesis, I conceive this current project as an exercise in the philosophy of mind and action as much as I see it as a project relevant to epistemology and ethics. Also, this relatively new interest in the philosophy of mind extends beyond its relation to the ethics of belief. Just one other way this interest has expressed itself under Lucy's influence, is its cumulation in a side-project – which perhaps I spent too much time on this academic year! – on emotion and the religious life, 'religious shame'. This too Lucy has helped with. Not only this, but Lucy readily agreed to take me on as a supervisee at a difficult and busy time for her. I'm extremely grateful to her for this. In short, Lucy has greatly impacted me as a philosophy student. I thank you, Lucy.

The virtues of John Hyman, the philosopher, are well known. Anyone slightly familiar with his more recent work will easily recognise his influence on my thinking. So much so, that it seems almost foolish to rehearse how he has been influential in this note of thanks. The numerous references made to his work in this thesis are but a symptom of his influence on me; they are by no means its essence. Less well known, probably, are his virtues as a supervisor. His patience, understanding, and honesty were really important to me. And I know to cope with me as a supervisee, they are qualities needed in abundance. I thank you, John, for all your time and effort, and for a lot more than this too.

There are people who were once in my life, but for one reason or another, are no longer. I also owe and feel a deep gratitude to them for their support. I know who they are, and I thank them.

Aḥaron, aḥaron ḥaviv. Last, but not least, my parents. There have been times in the past years which have been very difficult. There was no guarantee that I'd get to this point of completing the MPhil – far from it. Through it all, my parents somehow put up with me. My studying could not have been possible without their moral and financial support, and most of all, their kindness. To say anymore would be to belittle what they have given me. 'Everything that can be put into words can be put clearly' (*TLP* 4.116). But I'm not sure if I can put clearly my gratitude. Certainly not in a way that would give full expression to it. Like the Psalmist who says, and Maimonides' *via negativa* which holds, 'to You [God], silence is praise'; to you, Mum and Dad, my silence is praise. I will, however, say thank you. I know this is no recompense, but Mum and Dad, with love, this is for you.

Contents

<i>Title page</i>	1
<i>Declaration</i>	2
<i>Epigraph</i>	3
<i>Abstract</i>	4
<i>Impact statement</i>	5
<i>Acknowledgements</i>	6
<i>Contents</i>	8
Introduction	10
Part 1: Psychological, Conceptual, and Normative Doxastic Involuntarism	12
1.1. 3 kinds of doxastic involuntarism in Aristotle	12
Part 2: A Discourse on Voluntariness	18
2.1. Voluntary action	18
2.1.1. ‘Volitional’, ‘willed’, or ‘conative’ voluntariness	20
2.1.2. Ryle’s objection	25
2.1.3. Who wears the trousers?	30
2.2. The ignorance condition	34
2.2.1. Error and ignorance	39
2.3. The compulsion condition	42
2.3.1. Compulsion, choice, and evaluation	47
2.4. Voluntariness and the will	50
2.5. Voluntariness and control	51
2.6. Voluntariness and intention	53
2.7. Voluntary agency and voluntary passivity	53
2.8. Voluntariness’ minimal condition	54
2.9. Conclusion	55
Part 3: Doxastic Voluntarism	56
3.1. Doxastic (in)voluntarism and (in)voluntary belief	56
3.2. The doxastic ignorance condition	57
3.3. The doxastic compulsion condition: On being ‘compelled to believe’	65
3.3.1. Normative doxastic involuntarism	66

3.3.2. Conceptual and psychological doxastic involuntarism	67
3.3.3. Voluntarism without the will: Un-willing beliefs	70
3.3.4. Evidentialism and compulsion again	71
3.4. Acts and states – agency, patiency	75
3.5. An objection – wanting, wishing, and ‘outward criteria’	75
3.6. Conclusion	77
Conclusion	79
<i>Bibliography</i>	80

Introduction

The aim of this essay is to motivate the thesis that doxastic voluntarism is *routinely* true. This claim is anathema to most epistemologists. Even voluntarists, I presume, would recoil from such bold declarations. Nevertheless, I aim to make it plausible, and I do so in a distinctive manner. This is because in this discussion, I start from the ground-up, as it were. A great deal of the literature involved in the controversy between doxastic voluntarists and involuntarists consists in arguments about our ability to believe ‘at will’, to ‘control’ our beliefs, and sometimes even to ‘intend’ our beliefs. There are certainly important discussions to be had bearing those concepts. But it had been a cause of growing frustration to me that little attention had been paid – albeit with some exception – to the *voluntariness* of doxastic involuntarism/voluntarism. If these epistemological positions are about anything, surely they are about the possibility of *voluntary* belief. Encumbered by this frustration and armed with the determination to rectify the philosophical landscape, I came to work on this thesis.

The thesis has its form in three Parts. Part 1 is by way of a primer to doxastic involuntarism. I present three different kinds of involuntarism. I do this by closely reading a compact but rich argument in Aristotle’s *De Anima*. Part 1’s dialectical function is purely preparatory. By understanding the terrain of doxastic involuntarism, readers will be better equipped to follow my use of the three modes of involuntarism individuated here, when I rely on their formulation later in the essay.

Part 2 begins with a central problem in the philosophy of action: how to distinguish between mere movement and action attributable to agents. I elucidate how attempts to address this problem created the conditions for philosophical confusion. By positing that action is a complex whole consisting in physical and mental elements, voluntary action had the misfortune to be seriously misconstrued. Voluntary action – at least *bodily* voluntary action – came to be understood as physical movements precipitated by acts of will, volition, or conations. The gravamen against this view of action is that it conflates two distinct axes of human conduct: the axis of *agency* which is to do with causing change, and the axis of *ethics* which is a distinctively evaluative affair. ‘Conflate’ is the important word for these axes can obviously *combine*. The problem occurs when voluntariness is *expropriated* or *subsumed* by the agential dimension. This occludes a proper grasp of both agency *simpliciter* and voluntariness,

on their own terms, and errors proliferate as a result, including within the doxastic voluntarist–involuntarist debate.

After unpacking this confusion, I consider Ryle’s objection to these volitional theories of action. I argue the objection isn’t fatal, but nonetheless, theories of action must not conflate voluntariness and agency as a basic criterion of adequacy.

The rest of Part 2 is devoted to defending a broadly Aristotelian account of voluntariness, as outlined by Hyman. Voluntariness is an evaluative quality attributable to *doxastic* and *evaluative* creatures. This only becomes clear once I have defended the idea that voluntariness is defined by double-negation. That is, where there is ignorance or compulsion, there is no voluntariness since these things typically, though not invariably, *exculpate* people from guilt and blame. And because ignorance and compulsion relates to creatures’ *believing* nature and a capacity to *desire*, only creatures which possess these features are eligible candidates for ascriptions of voluntariness and involuntariness. Throughout this Part, I argue and explain for a particular view about the conditions which cancel voluntariness. The ignorance condition is defined in 2.2, and the compulsion condition is defined in 2.3. I close Part 2 by considering the relation between voluntariness so argued for, and the will, control, intention, agency, and passivity.

In Part 3, I apply my findings from Part 2 to the doxastic case. I argue that doxastic voluntarism is *routinely* true. This becomes clear once we un-will notions of will from doxastic voluntarism/involuntarism. Doxastic voluntarism *isn’t* a thesis about whether we can believe at will, or about whether we can exert our will in such a way to control our beliefs. Rather, it is a thesis about whether believers are *prima facie* appropriate targets of doxastic evaluation. I present the doxastic conditions for doxastic involuntarism. I argue that these conditions are generally not satisfied. Therefore, the consequences of undoing the damage done by conflating agential action with voluntariness have substantial repercussions in the epistemological sphere. By freeing voluntarism from burdensome notions of believing at will, by *un-willing beliefs*, doxastic voluntarism becomes far more attractive and far less resistible.

Part 1: Psychological, Conceptual, and Normative Doxastic Involuntarism

1.1. 3 kinds of doxastic involuntarism in Aristotle

It is a fraught topic how best to understand Aristotle and his predecessors' conceptions of epistemic concepts like knowledge and belief. In *Meno*, Plato suggests via Socrates that knowledge and belief are part and parcel of the same epistemic genus. They are made distinct by the fact that belief, in contrast to knowledge, is an imperfect, impoverished grasp of the same truths as knowledge (Guthrie, 1956, p.113). Aristotle echoes this when he compares someone who believes to someone who is sick. The believer is unhealthy with respect to truth (*Metaph. Γ 4*, 1008b27-31). Further, belief differs from knowledge in terms of how secure they are in the subject. The one who has knowledge of p has p stably, whereas one who believes that p has a much looser relation to p (*Meno*, 98a).

But famously in *Republic V*, we encounter Plato's argument to the effect that belief and knowledge are two separate faculties extending over different objects (477a-478e). In this 'two worlds' or 'distinct objects' epistemology, one can know about Forms but not about sensibles. One can have beliefs about sensibles but not about Forms (Fine, 1978, p.121). They are two distinct faculties (*dunameis*), operating over two distinct sets of objects, accomplishing distinct ends.¹ This reading, which was once dominant, has undergone much criticism in recent decades to the point where the distinct objects interpretation of *Republic V* is considered by some to be 'outrageous' (Baltzy, 1997, p.240).² Our particular concern is how best to treat Aristotle's claim in *De Anima (DA)* 3.3. Aristotle argues that *hupolēpsis* is not the same kind of mental thing as *phantasia*. Here is the argument:

That [imagination] is not the same kind of thinking as *hupolēpsis* is clear. For this affection [imagination] is up to us whenever we wish (for we can put something before our eyes, just like those who place and form an image in mnemonics), but *doxazein* [the verbal form of *doxa*] is not up to us, because it is necessary either to be true or to be false. Further, whenever we

¹ This is Socrates' litmus test for individuating differing faculties or powers (*Republic* 477d).

² Moss' spirited defence of the distinct objects reading hinges on the observation that recent scholars of Plato have been guilty of eisegesis (2021, p.42). See Moss (*ibid.*, pp.20f) for a collation of criticisms raised against the distinct objects reading.

doxazein something terrible or fearful, we are immediately affected, and similarly with something encouraging; but with respect to imagination we are in the same condition as if we were observing terrible or encouraging things in a painting. And there are differentiae of *hupolēpsis* itself: knowledge, *doxa*, practical wisdom and the opposites of these; of these differentiae I must speak elsewhere. (427b16-27)

How ought we translate key epistemic terms in this passage? I have encountered three different renderings of *hupolēpsis*: (a) ‘conceiving’ in Shields’ translation (2016); (b) ‘supposition’ in McCready-Flora (2013, p.85); and (c) ‘belief’ in Moss and Schwab (2019). *Hupolēpsis* is no filler-word. It is the very thing that Aristotle is in the business of showing is different to *phantasia*, imagination or appearance. So, the translation choice has substantial repercussions for Aristotle’s argument.

This translation is copied from Moss and Schwab (2019, p.17). They aim to provide excellent evidence that *doxa* ought not be translated as ‘belief’, as that term is ordinarily understood in contemporary epistemology (ibid., p.14). Aristotle imposes order on the conceptual baggage he inherited from Plato by providing a consistent framework to use *doxa*, as well as a generalised theory of *hupolēpsis*. Thus, for Aristotle, *hupolēpsis* is used as a term for ‘the genus of other cognitive attitudes and that it plays this role because it is the attitude of taking-to-be-true as such’ (ibid.).

In following Moss and Schwab, we are taking *hupolēpsis* to mean ‘belief’. That is, a term to denote a cognitive attitude of taking-to-be-true something or other. Aristotle’s argument would otherwise lose some of its pertinence to our discussion which, for Moss and Schwab, it directly possesses. The first reason Aristotle gives to support his conclusion that imagination is not the same thing as belief, is that imagination is a mental act which is ‘up to us whenever we wish’ whereas *doxa* is ‘not up to us’. Why in this premise does Aristotle use *doxa* and not *hupolēpsis* if it is the latter that the argument primarily aspires to distinguish from imagination? It must be that there is some property of *doxa* which is also shared by belief viz. a cognitive attitude such that it takes something to be true (ibid.). Consequently, whatever we say about *doxa* in this passage we can say about belief.³ Before inferring that Aristotle is a doxastic involuntarist, it does well to think about what the correct reading of ‘up to us’ is.

³ And if Moss and Schwab are wrong and *doxa* ought to be translated as belief, then what Aristotle says in *DA* 3.3 about *doxa* just is what he says about belief.

The reason, says Aristotle, why *doxa* isn't up to us is because all *doxa* necessarily have truth-values. This reason is strange because it is surely possible that every imagined item can either be true or false. We can ask of every imagined mental representation if the contents of the affection correspond to reality. I can imagine Boris Johnson is the PM of the United Kingdom (it turns out he is) but this is not the same thing as believing the same content. So, it is not up to me to settle the truth-value of my imaginings in just the same way it isn't up to me to settle the truth-value of my *doxa*. Moreover, Aristotle himself attributes truth-value to imaginings e.g., *DA* (428a11). Hence, the precise meaning of the not-up-to-usness of *doxa* and its dialectical function is obscure.

The up-to-usness claim and its negation appears in the *Nicomachean Ethics* (*NE*) in Aristotle's discussion of moral responsibility and voluntary action.⁴ When the source of the action or its moving principle is located within the agent herself, we have good grounds – though not sufficient reason – to individuate that action as voluntary. This is how people can become ill or vicious voluntarily (*NE* 1114a10-22). An ill person cannot, *just like that*, make herself better. But this restriction on her power doesn't vitiate her voluntariness in *becoming* ill if she became ill by indolence (*ibid.*). This is because the power to not become ill was originally in her, so it was up-to-her whether to become ill or not. Thus, in *NE*, whether something is up-to-us depends on that thing's *source* or *moving principle*. If the source/moving principle of the action or state of affairs is found to be within the agent herself, it was up-to-her to do that thing. This makes the up-to-usness claim sound like it is a claim about the sort of *capacities* agents have or lack. Is this the best way to understand Aristotle's use of the same words in his argument in *DA* – that beliefs are not-up-to-us because their moving principle is not wholly within us? That we lack a capacity to believe in a way totally up-to-us, but we have a capacity for imagination which is fully up-to-us.

This is one way to ascribe doxastic involuntarism to Aristotle. Human beings are in want of the psychological wherewithal to believe in an up-to-us way, because we don't have the capacity to believe up-to-us.

⁴ Usefully mentioned by McCready-Fora (p.86).

Psychological doxastic involuntarism: ordinary human beings (creatures like us) lack the psychological capacities to believe voluntarily.

However, there is no mention of source, of *archē* in our passage. Therefore, it's unlikely that Aristotle is claiming something about a particular capacity for action (believing something) which we lack.

There are other ways philosophers can fall under the banner of doxastic involuntarism. One might argue that there is something about the very concept of belief which precludes believing at will. Call this, *conceptual doxastic involuntarism*. The truth of doxastic involuntarism is no contingent matter, like it is for psychological doxastic involuntarism. It is a necessary truth falling out from the nature of belief itself.

Conceptual doxastic involuntarism: the nature of belief makes doxastic voluntarism a conceptual impossibility.

It is perhaps natural to impute conceptual doxastic involuntarism to *DA* 3.3. Recall that the argument contrasts *doxa* and imagination on account that the former necessarily carries a truth-value whereas imagination doesn't. Somehow, this makes imagination up-to-us and belief not-up-to-us. The fact that Aristotle thinks imagination is also suitable for truth and falsity creates the following need: to interpret our passage in such a way that satiates Aristotle's distinction without exiling truth and falsity from the purview of imagination. This can be achieved by advancing an assertoric reading of the argument. As Wedin says, Aristotle is stressing the...

‘...logical point that the truth or falsity of something is crucial to whether we believe it in the sense that if a believes p, then a believes p is true. No one believes known falsehoods. It is just in this sense that belief is not entirely up to us...[Imagination] is essentially a [re]⁵presentational capability that in its own right asserts nothing about the way things are.’
(1988, p.76)

Whenever we believe, we inhabit a state which asserts something about the world. Belief's representational contents aren't neutral with respect to truth. The representational content in

⁵ Wedin's brackets.

belief *says* something about the world. Moreover, this is part of the conceptual constitution of belief. Beliefs always and invariably assert that something is the case. As a result, truth and falsity necessarily attach to beliefs. This is different to imagination. Whilst imagination represents to the subject certain images, its images don't assert anything about the world. Nevertheless, imagination can still be true or false in the sense that an 'image can be true *of* or false *of* something' (ibid.). I understand Wedin to mean that mental images which present themselves in episodes of imagination don't represent propositions, but they can be true or false with respect to that which they are an imaginary representation of. My imaginary representation of graduating with a PhD in Philosophy is true *of* that scene, so long as it represents the aspects and dimensions of that scene in a way that honours the verisimilitude of a *real perception* of that scene under ordinary conditions (ibid., p.79).⁶ Thus, Wedin's assertoric interpretation of the argument in *DA* 3.3 – endorsed by Moss and Schwab (p.18) – paints Aristotle in the light of conceptual doxastic involuntarism.

One might also want to propose a *normative* kind of involuntarism. This iteration makes no claims about what is conceptually necessary nor about human psychological capacities. At root, it is a claim about by what standard one *ought* to believe something (or disbelieve or suspend judgment). Normative doxastic involuntarism bypasses complications about belief *qua* concept and human psychology to proffer an ethics of belief. That is, it asserts that there is some standard by which a belief can be evaluated as in some sense 'good' or 'appropriate' and their opposites. One can reasonably understand normative doxastic involuntarism to be primarily concerned with delineating the scope of our doxastic *freedom*. By claiming that there is some standard to which belief is subject, it suggests that there are ways of believing which are doxastically out of bounds if we want to comply with doxastic norms. Because it says something about what sorts of doxastic practice are epistemically illegitimate and legitimate, it delimits the extent of our doxastic (normative) freedom.

Normative doxastic involuntarism: belief is subject to some [unstipulated] standard, and one's freedom to believe is constrained by that standard.

⁶ Admittedly, the chosen example is tricky because it is a prospective imagination of a scene which (hopefully) will be perceived, and it contains so many elements. For ease, imagine any ordinary 'middle-sized dry goods', like a chair or table. The mental image in the imagination is true of the chair so long as it represents the chair accurately according to a veridical perception of the chair under usual conditions.

McCready-Flora holds that *DA* 3.3 is best construed as positing normative doxastic involuntarism. The fact that belief, unlike imagination, isn't up-to-us has nothing to do with our incapacity to believe at will; it is about the norms which govern belief. To support his reading, McCready-Flora cites from *Eudemian Ethics* (*EE*) 7.1 which is about the nature of friendship. There, Aristotle distinguishes between what governs behaviour amongst friends and what governs behaviour amongst others (strangers).

Private right conduct towards our friends depends on ourselves alone, while our conduct towards others is determined by law, and is not up to us. (1235a2-3).⁷

Aristotle isn't saying that an agent's interactions with others isn't up-to-her because the moving principle or source of those interactions isn't to be found within her (McCready-Flora, p.89). The point is that agents cannot do as 'they jolly well like' with others because that arena of human affairs is regulated by law. It's not-up-to-us because given certain stimuli – and the assumption that the law works – the law will react and intervene to manage behaviour amongst its citizens. Whereas right conduct amongst friends is secured by the very fabric of friendship.⁸ No intervening power is required. 'We form these relationships and take these [private, right actions] just because we want to, which makes them up to us' (*ibid.*, p.90).

Thus, when Aristotle tells us that belief is not-up-to-us, he means that belief formation is governed by a norm. And when Aristotle instructs that imagination *is* up-to-us, he means that contra belief, imagination isn't subject to norms. We are free to imagine whatever we like, but we are not free to believe whatever we like.

In conclusion, our reading of *DA* 3.3. yields three different kinds of doxastic involuntarisms: psychological, conceptual, and normative doxastic involuntarism. Aristotle does *not* argue for psychological involuntarism in *DA* 3.3., though one can reasonably interpret him to be arguing for conceptual or normative involuntarism. Acquaintance with the variety of doxastic involuntarisms will be profitable as we proceed to investigate the nature of voluntary action and the relation between voluntariness and belief.

⁷ To keep the 'up to us' phrase consistent, I have amended Kenny's translation.

⁸ See McCready-Flora (pp.89f) for an explanation of this contention.

Part 2: A Discourse on Voluntariness⁹

Part 2 aims to understand the nature of the voluntary. It opens by explaining how a central problem in the philosophy of action – how to distinguish between mere movement and action attributable to agents – encouraged erroneous thinking about voluntariness. This is because theories of action relying on *acts of will* and the like, conflated action as such with voluntariness (2.1-2.1.2.). 2.1.1. explores the history of this error and 2.1.2. considers whether Ryle’s objection is as fatal to volitional theories of action as it is sometimes thought to be.

I continue by arguing for and defending a broadly Aristotelian account of voluntariness, one which Hyman (2015) has (with some divergence) presented. At root, voluntariness is an ethical concept defined by double negation: voluntariness is incompatible with ignorance and compulsion. I explore these conditions of involuntariness (2.1.3.-2.3.1.). The end of this Part is devoted to investigating the relation between voluntariness and the will, control, intention, agency, passivity (2.4.-2.7.). Finally, I argue that our findings also reveal a positive minimal condition for voluntariness. Creatures to whom ascriptions of voluntariness and involuntariness can legitimately be applied are *doxastic* and *evaluative* creatures.

2.1 *Voluntary action*

Any analysis of action worth its salt must offer some way to distinguish between movements that are *done* and movements that *merely happen*. In the English language, our use of terms to refer to actions is not always consistent. It also does not always cut across philosophical differences between different kinds of movement. ‘Act’ and ‘action’ might be colloquially used to mean the same thing. But a little reflection reveals that not every movement is an action, and not every action is an act. For example, the raising of my hand is a movement but it may not count as an action if it was raised in the midst of a body spasm. If some criteria are satisfied, the raising of my hand could very well be an action e.g., if I did it on my own bidding. But as well as the action-type of hand raising, there is also the act of getting someone’s attention *by* raising one’s hand. When I’m in the crowd of the Pyramid Stage at Glastonbury, I can easily perform an action of raising my hand. But that action doesn’t count as the finished act of getting someone specific to notice me on account of my hand raising if

⁹ In this section, all references to Hyman are to Hyman (2015) unless otherwise stated.

that specific person didn't notice me on account of my hand raising.¹⁰ So, we have a taxonomy which distinguishes between different classes of movement:

Mere movements: the physical exertion of my muscles and/or my bodily frame which results in a change of my body's position from P1 to P2.

Actions: the physical exertion of my muscles and/or my bodily frame which results in a change of my body's position from P1 to P2 *in virtue* of some special relation which connects me *qua agent*, to the exertions of my body in the right way.

Acts: actions which are seen through to the end such that typically, although not always, the agent's purpose for which the action was done is satisfied.¹¹ Acts express *that which is done* when the action is carried out.

The distinction between mere movements and actions can also be cashed-out in terms of intransitive and transitive movement (Hornsby, 1980, pp.2f). English often carries ambiguity in verb-use. When, for instance, noun-phrases are used referring to the 'melting of the chocolate' or the 'growing of the lettuce' it isn't always plain whether what is being referred to is the transitive melting/growing of the chocolate/lettuce – which is a person doing something to the chocolate/lettuce – or the intransitive melting of the chocolate/lettuce – in which the chocolate/lettuce is the object undergoing a change in status (ibid.).¹² To signify this difference, Hornsby uses [verb]_T to indicate transitive use of verbs and [verb]_I to indicate their intransitive use. The verb 'to move' is similarly subject to ambiguity. The 'moving of *x*' can either be the movement_T of *x*, or *x*'s movement_I. The latter are, for Hornsby, the results or effects of some doing, so bodily movements_I cannot be actions. But movements_T arise because an agent has caused them or determined to bring them about. Therefore, if actions are bodily movements, they are bodily movements_T.¹³

¹⁰ It will of course, still be an *act* of raising my hand. But that act isn't sufficient to count as an act of *signalling* or *getting noticed* by raising my hand.

¹¹ But they are not a subclass of actions (White, 1968, p.2). There are *acts* of mercy but no corresponding actions of mercy (ibid.).

¹² See Steward (2012, pp.33f) for a useful explanation of Hornsby's distinction.

¹³ Interestingly, not every language suffers from this ambiguity. In Biblical and Modern Hebrew, the *hiphil* verb-stem functions to express causative action. And this, although perhaps not exactly, nicely enough maps onto the in/transitive distinction which English speakers need to work out from context. That lettuce grows because *I* grow it, is derivable in Hebrew by use of the causative 'to grow' viz., *magdil*. (Although it is also true that Hebrew possesses other verbs to denote the action of someone growing agricultural or horticultural objects e.g., *litzmoach* (inf. to grow; as in to cause something to sprout) or *lifroach* (inf. to grow; as in, to cause something to flower)

To make the above categories clearer, consider another example. Unfortunately, we live in an unpredictable world ridden with tragedy. Horace is attending a roof-top party in a swanky London hotel. The atmosphere, aided by abundant quantities of alcohol, is lively. Another guest, Ovid, who is dancing, wildly bumps into Horace with sufficient force to knock him over the edge of the roof. Luckily for Horace, but extremely unluckily for the hotel's doorman, the latter was standing in just the right place to break Horace's fall resulting in the doorman's untimely death. Horace's mere bodily movement brought about the death of another.

Horace, believing under false pretences that his being knocked-off the hotel's roof was a deliberate attempt on his life, aims for recompense: he intends to murder Ovid. Horace shooting Ovid is an action with the view to murdering Ovid. The shooting itself is an *action*, but so long as Ovid remains alive, Horace hasn't committed the act of murder. This is because murder is the unlawful killing of another human, so murder requires someone's death. If Ovid survives the shooting, at least for a while, we cannot say Horace has committed/carried out an act of murder, though we can say he intended to murder. If Ovid succumbs, Horace has committed the *act* of murder.

2.1.1. 'Volitional', 'willed', or 'conative' voluntariness

If it is presumed that the central task for the philosophy of action is to account for the distinction between actions and mere movements, then, as Hornsby rightly points out, it is only 'natural' to search for answers in a movement's *causal* history (p.50). In turn, the distinction between what movement is caused by some sort – the *right* sort – of psychological event in the agent and what is caused by something 'outside one's control' marks the distinction between action and mere movement (ibid.). Alternatively, if these events are not causal, they are nevertheless presumed to be ineliminable to understanding an action's

which perhaps renders one being '*magdil* lettuce' an awkward construction). Moreover, some verbs directly map onto the in/transitive distinction: *lilmod* (inf. to learn) is to *lelamed* (inf. to teach); *le'echol* (inf. to eat) is to *le'ha'achil* (inf. to feed i.e., to cause something else to eat). See Hyman (p.36).

The following may also be of interest to the reader. Hebrew has no grammatical means to distinguish between the progressive and perfective aspects, where English can. Thus, it can get tricky figuring out whether an *act* or *action* occurred, in the sense prescribed by the definitions above. In English, I can express my unfinished act of building my house ("when I was building my house") and my finished act of building my house ("when I built my house"). There is no means of expressing this difference in Hebrew other than using extra words to describe at what point along the process I was up to with my building.

provenance. Causal or something less than causal, these mental antecedents become intertwined with what it is to act (or to make an action).

Another related reason for the search for mental causes or provenances to action, is the fact that desires and beliefs are not sufficient to generate action. I can undergo desperate cravings for cigarettes yet not smoke. I can be genuinely repulsed by the idea of my smoking, believing it to cause cancer, and yet choose to smoke. Hence a mediating operation of mind is postulated 'by means of which a mind gets its ideas translated into action' (Ryle, 2000, p.62). So, desires and beliefs are non-executive. They alone don't get the agent to act. Thus, the concept of action is to be understood in terms of a movement's mental history: these are antecedent mental events which are causally efficacious and/or executive in getting the agent to act. What sort of antecedent mental event might these be? They are, *volitions, willings, acts of will, decisions, intentions*, or even *thoughts* (ibid., p.47). In Hornsby's quasi-technical use of the term, they are grouped under the generic name of *conations* (ibid.f). These volitional events of mind are a prerequisite for action. Without their presence in a motion's causal sequence, an agent may move but she certainly won't act.

I shall not arbitrate whether theories of action which postulate mental events of these types to make sense of action are good on their own terms. Although, they are subject to considerable criticism.¹⁴ My immediate concern is the role voluntariness plays in these theories. This is because an action characterised as brought about by a conation came to be confused with an action that was voluntary. In this way voluntariness morphed into a different beast entirely, and became a backhand, confused way to speak about *both* the agential causes of an action *and* the type of human conduct which can appropriately be subject to moral predication (Hyman, p.5). Needless to say, any theory of action which conflates and thereby corrupts the notion of voluntariness with conations is *pro tanto* impaired to the extent it is guilty of this conflation and corruption. But that is not sufficient to show that some mark of the mental isn't necessary to demarcate action from mere movement, and that mark may yet be a causal one. But to repeat, I shall not arbitrate on this.

¹⁴ For example, see Frankfurt (1998b, pp.69-71) for his criticism of causal theories of action. See O'Brien (2007, pp.129ff, especially pp.129-134) for criticism of another salient feature of these theories: that they are *composite complexes* comprising mental and physical elements.

In *The Passions of the Soul*, Descartes writes that there are two principal kinds of thoughts: actions of the soul and passions (§17). Actions of the soul are termed ‘volitions’ (ibid.). Some volitions terminate in the soul for instance, when ‘we will to love God’ and some terminate in our body, ‘as when our merely willing to walk has the consequence that our legs move and we walk’ (ibid., §18) These actions of mind occur in ‘the will’. So actions of mind, volitions, are what generates bodily action and they occur in the will. However, in *Principles*, we encounter an important conflation between what it is for an agent to *act* (or to do actions), and what it is for an agent to be a legitimate target for praise (§37). Especially noteworthy, is Descartes slipping from talking about *not* doing something out of necessity to voluntary action (ibid.). Therefore, already in Descartes, two different axes of human conduct are combined together in one. The axis of agency – what it is to act – is mixed with the axis of ethical evaluation – what it is to act voluntarily such that one is properly subject to praise.

The distinction is further blurred in Hobbes’ *Leviathan*, where he explicitly talks in terms of *acts* of will ‘(not the faculty)’ (1.6.53). These acts of will are necessary components of all animal motion. Animal motion opposes vital motion. Vital motions are the sort of movements that all life forms display which are congruent to the vital sustenance of that life form. Movements like blood flowing through one’s veins, breathing, digestions, and excretion (1.6.1). Animal motion is *voluntary* motion. It needs the imaginative powers of the agent to get started; to get the body to move as imagined by our minds. Consequently, voluntariness is incorporated into Hobbes’ conception of agency. All motion like ‘*going, speaking* and the like’ have their source in the agent’s imagination, and this makes the motion voluntary. The confusion culminates in Hobbes’ idea that an act of will is the final stage in the process of deliberation, since this makes it the case that ‘a *voluntary act* is that, which proceedeth from the *will*, and no other’ (1.6.53). Thus, not only is voluntariness conflated with acting by dint of the will, it turns out that for Hobbes, every instance of bodily action is caused by a prior mental act.

Pink explains that by Hobbes’ own lights, voluntary action has its origin in involuntary motivations (2009, p.106). A motivation is a pro-attitude toward carrying out certain actions and outcomes such that they lean an agent to ‘favour’ the objects of those attitudes (ibid., p.103). Crucially, justifications for actions apply equally to these motivations. The justification will cite the reason grounding the pro-attitude. For instance, my motivation to take a shower is justified by what makes the content of that attitude desirable: the fact that it will cool me on this hot day. If motivations do not proceed from the will, then it is true that Hobbes holds that

motivations are involuntary. If Hobbes' dictum in (1.6.53) is meant as a catch-all pronouncement to apply to all action such that the only action there can be is *voluntary* action, then how are we to understand Hobbes' talk of *acts* of will? As Pink observes, in *Of Liberty and Necessity*, Hobbes writes that he has the liberty to *do* if he *wills*, 'but to say I can will if I will, I take to be an absurd speech' (§3). That is to say that episodes of willing don't proceed from the will because it is impossible to will oneself to will something. In which case, Hobbes' acts of will don't seem like acts at all, but passive occurrences. Thus, we have a textual difficulty. Hobbes says 'acts of will'; but he also argues that action is just the voluntary locomotive effects of imagination together with the will, a position which precludes the possibility of genuine acts of will. Perhaps Hobbes is speaking metaphorically in using the term 'acts of will'. Nevertheless, it is important to acknowledge that the existence of this textual difficulty and the way we choose to treat it, is independent from the objectionable fact that the kernel of Hobbes' theory of action blends action *simpliciter* with voluntariness.

Locke is creditworthy for his acute insight into the nature of voluntariness.¹⁵ Voluntariness, explains Locke, opposes *involuntariness*, not necessity (*Essay*, 2.21.11). He is, out of the thinkers surveyed so far in this section, the most sensitive to the difference between voluntariness and freedom, between evaluative notions and notions of power. Locke clearly accepts that unlike liberty, voluntariness is tied to preference and choice (2.21.10) (and so, is in this sense evaluative) and effectively dispatches the claim that freedom is anything but a power or ability attributable to agents or substances (2.21.16). This is why to ask if freedom is free is just as nonsensical to ask whether riches be rich. 'Rich' is an adjective ascribable to items such as persons, resources, and institutions. Being a 'denomination from the possession of riches' it is only accurately ascribed to such a thing if that thing possesses riches. Also, the question of the richness of riches can never arise (*ibid.*). Similarly, freedom is something that can only be predicated of agents and substances since it is they which harbour powers and abilities. One power cannot be said to be in possession of a power, let alone the very same power.

Yet, he is not totally immune from the charge that he has attached volitional baggage to his account of voluntariness and as a result, produced confusion. Locke's picture of the will is as a power of *preference* (2.21.5). It is a power of the mind which instructs the subject upon ideas

¹⁵ See Part 2.3.

being considered viz., ideas of possible actions, outcomes, events, etc. It instructs whether to forbear the motion preferring the body's rest instead, continue or end actions of our minds and motions of our body, or order the body to move (ibid.). This is the *power* of the will. Instances in which this power has been exercised are called *volitions* or *willings* (ibid.). The commission or omission of an act, as a consequence of a volition, is what makes voluntary action (ibid.). 'And whatsoever action is performed without such a thought of the mind [a volition] is called *involuntary*' (ibid.).

Perhaps two more illustrations of conations mixing with voluntariness will be useful. Reid indicates that he is also party to this confusion when he denies that the capacity to *undergo* change is a power. Criticising Locke's 'passive powers', Reid writes that the phrase is a 'contradiction in terms' and that Locke 'seems to have been unlucky in inventing it' (2010, p.21). This is significant because it suggests, although does not decisively demonstrate, that for Reid, if all power is *active* power and its exertion is dependent upon a conation, then voluntariness cannot coexist with passivity. The reason for this is that passivity and voluntariness might be mutually exclusive *because* voluntariness is inherently something which only attaches to *activity*. If this is Reid's position on voluntariness, this *eo ipso* suggests that Reid confounds voluntariness with agency. That Reid does so confound is confirmed by his opposing voluntariness to necessity:

If, in any action, he had power to will what he did, or not to will it, in that action he is free.¹⁶ But if, in every *voluntary* action, the determination of his will be the necessary consequence of something involuntary in the state of his mind, or of something in his external circumstances, he is not free; he has not what I call the liberty of a moral agent, but is subject to *necessity*.¹⁷ (ibid., p.267)

Further evidence that Reid overlaps activity and voluntariness is yielded by his claim that it is probable that our concept of active power emanates from our voluntary exertions (p.204) and in his assertion that a crime must be something voluntary because 'if it be not voluntary, it is no deed of the man...' (p.224). Finally, similar to his philosophic predecessors, the exercise of an active power always follows a conation, a determination of the will (ibid., p.59).

¹⁶ Note Reid's rejection of the Hobbesian doctrine that one cannot determine one's will. For Reid, action is only free if it were possible for the actor to have willed otherwise (Alvarez, 2010, p.508).

¹⁷ My emphasis.

Lastly, Bentham acknowledges the words ‘voluntary’ and ‘involuntary’ are ‘extreme’ in their ambiguous signification (1789, 8.2, note a). And though he succeeds in identifying one of the conditions which *does* make an action involuntary viz., coercion (ibid.), he nevertheless stipulates that when involuntary acts are performed, ‘the will has no sort of share: such as the contraction of the heart and arteries’ (ibid., 8.5, note b). But, as we are soon to see, it is false to say that the will is destroyed or negated in all cases of involuntary action. Moreover, although in this note Bentham seems to use a sentence construction which permits the locution ‘involuntary acts’, it actually appears that he thinks there can be no such thing. In note b of 8.5, involuntariness counters intentional (so that involuntary ‘acts’ are *unintentional*) and Bentham also writes, ‘[i]f the act be not intentional in the first stage, it is no act of yours’ (8.5). Which is to say an agent’s action is marked by a prior conation.¹⁸

2.1.2. Ryle’s objection

Ryle’s regress objection is well known but still worth rehearsing. It is thought to be a decisive objection against this cluster of volitionist¹⁹ views. In this section I recapitulate the objection. Then, citing Hornsby, I argue that its vigour is limited. Afterwards, I reject Hyman’s response to Hornsby. This leaves open the possibility of conationist theories of action, provided it meets certain criteria. I close by reiterating reasons for abjuring conationist theories of action.

Ryle raises four objections, but we shall focus on the last (pp.65f). The problem is this: volitions or conations were postulated as effective causes of voluntary action. Thus, it turns out that to every instance of action the volitionist points to, she surreptitiously points to a conjunctive proposition consisting in two actions: (i) the bodily action φ and (ii) the mental act/conation which caused φ , φ_m . But if φ_m truly caused φ so that it can explain why φ is voluntary, φ_m must also be the effect of a voluntary cause. It better be, or else the voluntary nature of φ is threatened. Hence the volitionist not only needs to postulate (ii) in addition to (i), but she now also needs to explain an additional voluntary cause to explain (ii). This will be (iii) a prior voluntary volition φ_{m-1} . What explains (iii)? Well –

¹⁸ The development of 2.1.1. has been largely guided by Hyman’s excellent overview of the ‘modern theory of the will’ (pp.211-221).

¹⁹ I use ‘volitionism’ interchangeably with ‘conationism’ and other cognate words.

- (iv) a prior voluntary volition ϕ_{m-2} , of course! And to causally explain (iv) we require
(v) ϕ_{m-3} and so on *ad infinitum*.

Volitional accounts of action which speak through the language of voluntariness are surely untenable if they imply this Rylean infinite regress. They must either take on board the regress or accept the *absurdity* of calling ϕ voluntary despite its cause being an *involuntary* conation.

Ryle's regress is eminently dialectically forceful. Yet, I have long struggled with seeing its fatal quality – fatal to volitional theories of voluntary action, that is – it is reputed to have. In other words, it seems to me that the power of the argument has been overstated. The reason for my doubt is that Ryle's target is specifically the doctrine of volitions qua *causal* hypothesis (p.66). But we have examined various iterations of the modern theory of the will/the doctrine of volitions/volitional voluntariness (however you want to label it), and a causal ingredient is not common to all. Happily, Hornsby provides a formulation which gives expression to my doubt.

Hornsby distinguishes between three types of conations (p.48).

- 1) Conations are actions, but not causes of actions.
- 2) Conations are not actions, but they cause actions.
- 3) Conations are actions *and* causes of actions.

If a conationist holds (1), then the regress is blocked because there was never an imputation of a causal property to conations in the first place. As such, philosophers holding (1) aren't pushed to find causal explanations of the cause adduced to explain the advent of ϕ . If (2), then it is unclear why the conationist is under any pressure to provide a causal explanation of that which causes the voluntary action, for she can just deny that the conation itself requires a further cause in order to award the action the status of voluntary (*ibid.*, p.49). Albeit, one might still have to explain away a puzzle about how voluntary action can emerge from nonactive causes. Thus, the conationist hypothesis needs to be both causal and active – (3) – for the objection to pack its punch. Hobbes might subscribe to (3) depending on how we interpret his 'acts of will', and perhaps Reid does too in thinking humans have power over

their determinations of will (Reid, p.240). The absurdity that Ryle's regress latches onto, is the absurdity that in every one instance of action, there are actually two, and if two, then an endless sequence of different actions. This is only fatal to (3).

Hyman argues that this method of restricting Ryle's regress to type-(3) only works if the theory as a whole is taken to be about human agency rather than voluntary action (pp.21-24). Let us deal with conationism qua theory of human agency first. Descartes posited conations to explain the *motion* of our bodies. Descartes didn't posit them to explain when our actions are voluntary and when not (ibid., p.23). Seen in this light, conationism evades the regress. This is because actions have a variety of causes, and these causes can be active or passive (ibid.). For instance, a conation can arise in me to 'get me' to run 100m on account of my teammate passing me the baton in a relay race. My body's motion clearly has an active cause since the conation is the result of another agent's action. My body's motion can also arise from conations caused via passivity. When I leave a room because someone's presence upsets me, and this affective response caused the conation which aroused my bodily movement to exit the room, this conation was brought about by passive means – someone's presence upsetting me – but my subsequent movement is no less my movement than when I ran the 100m in the relay race. As Hyman writes, the 'variety of causes does not disqualify the volition [or conation] from making its effect attributable to my agency' (ibid.). The upshot is that specifically agential conationism is never pressed to explain the prior causes of action because the 'theory does *not* imply that whenever there is one action there are two' (ibid., p.24). There is no absurdity in thinking that conations have passive causes so there is no need to reach farther back for *active* explanations of conations. Neither is there any absurdity in thinking that an action stemming from conations that came about passively, and which I suffer, is *my* action. Thus, conationism as a theory of agency eschews Ryle's regress.²⁰

Consequently, Hyman argues that Hornsby's defence against Ryle's regress lacks potency precisely because she understands conationism to be a theory of agency. In other words, in

²⁰ One might protest that even if conations arise passively, they cannot be wholly passive events given their role in constituting that by which agents act. So, there still needs to be two actions for every action. However, this expostulation misconstrues the active potential of passive events. Assume that avalanches occur passively. That – the snow hurdling down the mountain – is no action, but it has tremendous potency to generate motion in other things, like a chalet. Analogously, a series of people coincidentally bumping into each other, accumulating more potential for motion as the series continues, is no action. Yet, someone in the series can utilise the momentum to help kick-start their run, which is an action. So, Cartesian-like conationism has no reason to concede that passive occurrences cannot be purely passive despite their function to produce agential action.

defending conationism against the regress by emphasising its qualities as a theory of agency, she is defending a fortress requiring no reinforcements. Hornsby, however, fails to notice that if conationism is a theory of *voluntary* action, her defence is helpless. This is because there *is* absurdity in thinking a voluntary action can arise from an *involuntary* cause. This is unlike action *simpliciter*, where no offence to reason is incurred by attributing an action to someone's agency even though it came about through a passive cause. Thus, where the defence *is* required, conationism qua theory of voluntary action, Hornsby fails to give it.

I object to these contentions of Hyman. Firstly, Hyman contrasts Kenny's reaction to Ryle's regress, with Hornsby's reaction (p.23). Hyman tells us that what grounds this difference in reaction is their view regarding about what exactly Ryle's target is. Kenny, in understanding conationism to be a theory about *voluntary* action thinks Ryle's objection 'decisive'²¹ viz., it decisively quashes the theory (ibid.). Hornsby, as we already know, thinks there is a defence in the offing. Only, Hyman tells us that Hornsby is defending a conationist theory of *agency* in general – this must be the reason for the contrast, otherwise it serves no purpose. So, we need to do two things here. (1) We need to determine whether Hornsby does indeed take herself to be defending conationism in its guise as a theory of agency. (2) And if this is *not* the case, we also need to determine whether conationism (of some type) can nevertheless be defended.

Apropos (1), Hyman's evidence that Hornsby is defending an agential conationism is insubstantial. It is true that Hornsby writes that the philosopher under Ryle's attack 'takes the occurrence of conations to distinguish between "mere bodily movements" and the voluntary movements that are actions' (pp.48f).²² But we are not forced to understand the 'that are actions' clause as predicating actions to a class of voluntary movements, which would make actions distinct from voluntary movements. Rather, the clause could be read by way of *identifying* actions with voluntary movements i.e., the type of theory which says that whenever we have one bona fide voluntary movement, we have one action. Thus, it isn't clear to me that Hornsby is attempting to defend conationism as a theory of agency as opposed to a theory of voluntary action. Now we must address (2). Is it true that as a theory of voluntary action, conationism cannot resist Ryle's dilemma?

²¹ Kenny's word.

²² Although in my version of the text, the last three words are not italicised as they are in Hyman's.

The way I want to answer this is by way of rejecting Hyman's distinction between two ways of reading conationism/volitionism. The putative distinction between conationism qua theory of agency vs. conationism qua theory of voluntary action, is artificial. This is because if Hyman's animadversions to the modern theory of the will are largely correct, it is doubtful that any advocate of the theory would have thought of their conationism as *either* a theory about agency *or* a theory about voluntary action. Conationism's distinctive mark is the failure to not only keep separate, but to positively amalgamate action with voluntariness. Thus, advocates of the empiricist position that conations cause or bring about action, ought not be distinguished between empiricist views about voluntary action. The two are combined to make *volitional*, *willed*, or *conative* voluntariness. And this is the foundational error. Hence, Hornsby's above defence cannot be a defence of one type of conationism as the expense of the other. If it is effective, it is effective for both (artificial) strains of conationism identified by Hyman. If it is ineffective, it is ineffective for both.

So, what about (2)? Can conationism defend itself against Ryle's dilemma? The answer is 'yes', but a qualified 'yes'. Hornsby is in the clear to develop a theory of action involving conations so long as voluntariness isn't confused with action.²³ Avoiding this confusion is a minimal criterion of adequacy. Furthermore, one would have to explicate what it is meant by conations' presence in a theory of action. For example, it cannot be some mysterious mental event which are never named in ordinary discourse. (This seemingly precludes volitions).²⁴ I submit – I think here in agreement with Hyman – that the real culprit which makes conationist accounts inviable, is their confounding of *voluntary* action with *what it is to act*. If action is understood as the sort of bodily motion which is voluntary because it proceeds from a conation, we fundamentally misunderstand both action and voluntariness on their own terms.

So, Melden is right when he argues that it is a hopeless endeavour to use voluntariness to mark the boundary between mere movement and action (1956, pp.524ff). Not, however, for the reason he thinks. Namely, that voluntariness is 'unilluminating' (ibid.). But because voluntariness is the wrong concept in the first instance to mobilise in pursuit of the desired

²³ For Hornsby, *tryings* – events which occur “right inside the body” prior to transitive muscle contraction and transitive bodily movements – seem to play the role of the conative act which, when they issue in transitive bodily movement, identify action (p.33).

²⁴ See Ryle (pp.63f).

distinction between movement and action.²⁵ There are various dimensions of human action. A concept which succeeds in elucidating something important in one of those dimensions will not necessarily succeed in elucidating a problem in another of those dimensions. The difference between action and mere movement concerns *agency as such* e.g., what it means to determine or cause change; what the difference is between activity and passivity. Voluntariness cannot clarify these issues. The difference between voluntary and involuntary action belongs to the *ethical* dimension of human action.

2.1.3. *Who wears the trousers?*

I have rejected the urge to understand voluntariness in terms of understanding the distinction between mere movement and action. Following Hyman, I argued that this hopelessly muddles things, conflating two different dimensions of human conduct: the dimension of what it is to simply act and the ethical dimension. This confusion is largely a result of volitional theories of action which in some shape or form, renders voluntary actions as the result of prior conations. Hence, voluntariness gets to be spoken about in the same breath as acts of volition, conation, and will.

Hyman's account of voluntariness, which I largely endorse, is broadly an Aristotelian one (*NE* 1109b30-1111b3). This is because like Aristotle, Hyman identifies two key conditions by which an action can be said to be *involuntary*. These are: (i) where the *arche* of the action is not in the agent herself and (ii) her being unaware of the particular circumstances of the action (*ibid.*, 1111a20-25).²⁶ Or, more succinctly: there are two conditions which undermine the voluntary nature of an act or action. They are:

- 1) Ignorance
- and
- 2) Compulsion.

²⁵ In an unabashed Wittgensteinian proposal, an action for Melden seems to be a bodily movement performed in a social context in which agents have the necessary hermeneutical tools to interpret the movement as *doing something*, and the movement was performed in order to arouse that interpretation. That is, it was done to *follow a rule* about how such movements are supposed to be interpreted as instances of doing ϕ (Melden, p.534). However, Melden's theory might tell us under what conditions an action can be said to occur, but it leaves us in want of an explanation as to why the agent did as she did (Davidson, 2001, pp.10f).

²⁶ Although in these lines Aristotle gives a positive definition of voluntariness rather than a definition by negation.

Thus, voluntariness is defined in terms of *double* negation (Hyman, 2016, p.694). To say something – be it an ‘act, omission, even, relation, condition, and so on’ – was done voluntarily, is just to say that it was *not* done by reason of ignorance *and* was *not* done by reason of compulsion (Hyman, p.77). An act was done *involuntarily* (in the sense of the negation of voluntarily) just in case it was done under the condition of ignorance or under the condition of compulsion. Before we tackle with more precision what (1) and (2) amount to, let me first argue, in defence of Hyman, why it is a good idea to account for voluntariness in terms of double negation.

When we ask, ‘What is voluntary action?’, we immediately encounter methodological trouble. The problem is that we all vaguely know what it is like to act voluntarily, but the experience seems to escape conceptualisation. For example, right now, the movement I’m making with my digits to exert pressure on the laptop’s keyboard, I know are done voluntarily. With little introspection, I know that I want to write these words, I’m trying to write these words, and I’m intending to do so. I’m confident that these strikes of my fingers are my choice, that there is no undue burden pushing me into this choice: there is no mental malaise making me do this, nor is there the bidding of another conscious agent making me do this. Moreover, I know what I’m doing in doing these movements.

All this goes to show that incipient notions of what it is to act voluntarily are not beyond the ordinary agent, even if its precise formulation is elusive. We are familiar with what it feels like, even if it cannot be so easily expressed. This is suggestive of a good methodology with which to tackle the question: to consider not what voluntary action *is*, but its opposite. One way to explain why voluntary action possesses this air of phenomenal familiarity and an aloofness from conceptual explication, is by observing that voluntary action is the default mode of human conduct. Voluntary action occurs when everything is normal, when everything is going right. One may draw an analogy with ‘feeling well’.²⁷ We all know what it is to feel well, but prod someone for an explanation and you will probably be unsatisfied. In response, you might receive a flummoxed look. Or, if someone were to really take the task on, she might give you a very long list of propositions detailing that in each part of her body which could be amiss, everything is *not* amiss. “My left foot is not in pain; my ankle isn’t swollen; my stomach

²⁷ To be read as ‘feeling *physically* well’.

doesn't ache" and so on. It appears far easier to describe an aberration from the normal, than the normal itself. I might know when I'm feeling well but I struggle to tell you why. To quote Wittgenstein, when writing voluntarily, 'one feels something, but could not possible analyse the feeling.' (1967, §586). In contrast, I know when I'm feeling unwell, and generally I can tell you exactly why – because I have a headache or an upset stomach.²⁸ (Or in a Wittgensteinian gloss: if it is a misnomer to say "I know I am unwell" 'except perhaps as a joke', I cannot doubt when I am feeling unwell since I *have* the very sensations which are the cause of my unwellness (2009, §246)). Consequently, although feeling physical well is normal, it is most efficiently defined in terms of deviations from the normal i.e., feeling unwell. I know that feeling well consists in the absence of any condition which would make me feel unwell.

There are other concepts which receive similar treatment. For example, Austin thinks the concept 'real' should be treated in like fashion.

It is usually thought, and I dare say usually rightly thought, that what one might call the affirmative use of a term is basic—that, to understand "x" we need to know what it is to be *x* or to be an *x*, and that knowing this apprises us of what it is *not* to be *x*, [or] not to be an *x*. But with "real" it is the negative use which wears the trousers. (Austin, 1962, p.70)

Real things are around us all the time. Eschewing metaphors, we can't help but live in reality. Real is the baseline. It is what is ordinary. Similar to feeling well, it is hard to get a conceptual grip on the real without having recourse to its negation, the unreal.²⁹ For aberrations from the ordinary jump out at us as phenomena requiring explanation. Once we are probed to explain the aberration, we are probed to think about *how* it deviates from the ordinary. Thus, real perceptual experiences can be distinguished from unreal, hallucinogenic perceptual experiences which present the same content to a subject, by identifying some feature of the hallucination which strays from the ordinary. In this case, the hallucinogenic perceptual experience is unreal because it is the result of some drug taken by the subject. Now we know that taking substances which stimulate hallucinations produce unreal perceptual experiences. From here we can infer at least this: that a real perceptual experience must *not* be caused by

²⁸ This is not a claim to the effect that people are *accurate* interoceptive reporters. It is just a claim that when people feel physically unwell, they generally can and do report a problem they feel in their body, even if their interoceptive faculties are poor. For bleak conclusions about whether current scientific data on interoception which relies on self-reports (questionnaires) converge on measuring the same thing, see Desmedt's et al. meta-analysis (2022, §5.2).

²⁹ Or 'not real' or 'anti-real'. Nothing is meant by my selecting 'unreal' other than the negation of 'real'.

hallucinogenic substances which interfere with one's perceptual faculties. This is to give a partial account of real perceptual experiences in terms of *unreal* perceptual experiences. We apprehend the former by first apprehending the latter. This is what Austin means when he says the negative use of 'real' 'wears the trousers' (*ibid.*). Real must be understood by referring to its rudimentary conceptual counterpart, the unreal.

The same can be said for voluntariness. *Involuntariness* 'wears the trousers', not voluntariness (Hyman, 2016, p.694). This makes voluntariness a concept of (double) negation. According to Hyman, to say something was done voluntarily, is just to say it was *not* done *involuntarily*. But why is it important for us to be cognisant of aberrations from voluntary conduct? Hyman argues that a sensitivity to actions and events which deviate from the ordinary i.e., voluntary conduct, have always played an integral role in our ethical lives. To notice this, dwell a moment on what the connection is between ignorance and compulsion (Hyman, p.77). Both factors are things which, when present, typically, though not always, exculpate (*ibid.*). They exculpate agents from being subject to negative moral evaluations. Further, their primary exculpatory force is legal. Ignorance and compulsion exculpate agents from guilt – not guilt the emotion, but legal guilt in the sense that ignorance and compulsion can block an accused from being liable to punishment by law (*ibid.*). Therefore, voluntariness is at root a legal, and more broadly, an ethical concept 'designed'³⁰ to help us assess agents' culpability.

Because legal systems and moral practices require methods of determining culpability, it was profitable to develop an alertness to cases in which human agents did not behave as they are ordinarily expected to behave *viz.*, voluntarily. When an agent is ignorant of her own doing or is compelled to do something, the conditions under which her action was performed are sufficiently deviant to raise the *prima facie* prospect of exculpating her.

This is why involuntariness 'wears the trousers' and voluntariness is a doubly negated concept. It is morally and legally advantageous to notice first and foremost digressions from normal human conduct because this permits the admissibility of moral and legal exculpations. As far as exculpations go, the legal and moral significance of regular, quotidian, and expected

³⁰ See Hyman (pp.77f) for a fascinating, if brief, overview of voluntariness' pervasive presence in a diverse plethora of legal systems and ethical codes in antiquity. The thrust of Hyman's point is that systems of law are almost inconceivable without these systems acknowledging exculpating factors by way of involuntariness (*ibid.*). This is intended to support the claim that voluntariness is a contrivance of moral imagination, specifically. That it was designed for ethical deployment.

human conduct is parasitic on its opposite, involuntariness. Voluntariness achieves its legal and moral function because if we know what it means for ϕ to have been done involuntarily, we know by extension that if ψ was done voluntarily, the same grounds for exculpation will *not* apply.

Following Hyman, I have stated that conditions (i) ignorance and (ii) compulsion negate voluntariness. I now turn to a more precise articulation of these conditions.

2.2. *The ignorance condition*

Ignorance is a way of being unaware. I don't want to claim that all cases of being ignorant are also always ways of being unaware, at least as we ordinarily use the word 'unaware'. For example, one can fail to be *comme il faut*, and so be unaware of social etiquette, without necessarily being ignorant in the way germane to voluntariness. Ignorance's etymology is derived from the Latin *in* (meaning: opposite) and *gnarus* (meaning: knowledgeable) (Peels, 2010, p.59). So, ignorance is an epistemic thing, and its etymology suggests that ignorance is a privation of knowledge. That is to say, the state of ignorance is the state of *not knowing* something. Not knowing what? As described by Peels (*ibid.*, pp.57f) there are different kinds of knowledge, and so different ways of knowing something. There is *experiential* knowledge which is the knowledge constituted by familiarity with an object's qualia e.g., knowing the taste of lemonade. There is *practical* knowledge which is the knowledge subjects have when they know *how* to do some activity e.g., riding a bike. Lastly, there is *propositional* knowledge, which is knowing the truth-value of p , knowing facts, knowing that something is or is not the case. Without committing myself to the reducibility of experiential and practical knowledge to propositional knowledge,³¹ I shall only focus on propositional knowledge because (a) even if there is not a relation of reduction between experiential and practical knowledge to propositional knowledge, there are still propositions which can be expressed about either experiential matters or practical matters, even if knowing *those* propositions isn't sufficient for experientially or practically knowing p ; and (b) to do otherwise would over complicate things.

³¹ If it is true that knowing-how to ϕ is reduceable to knowing a class of propositions about ϕ , failing to be *comme il faut* will count as ignorant unawareness.

The etymology and ordinary language use recommends the *standard view* of ignorance which is that ignorance about p is the privation of (propositional) knowledge about p (Le Morvan & Peels, 2016, pp.15f). On this view, ignorance is a complementary or contradictory of knowledge (ibid.). Hence, when S knows p , she is not ignorant of p . When S doesn't know q , she is ignorant of q . Therefore, exculpatory ignorance (EI) is:

EI: if S does ϕ not knowing some proposition relevant to ϕ , S is exculpated from blame for ϕ -ing.

Is this right?

This isn't even *ab initio* right. To see why, we need to make two further distinctions about ignorance. Firstly, Aristotle makes clear that there is a difference between acting *in* ignorance on the one hand and acting *by reason* of ignorance on the other (*NE* 1110b20-1111a1). A drunk man doesn't stop knowing things. Although, he is in a condition which impairs his thinking. His acts then, nefarious or virtuous, are performed in ignorance, which is to say, in a state of ignorance. However, someone who ϕ -s by reason of ignorance because he is (to use Aristotle's phrase) ignorant of the *particulars* in the case – e.g., he doesn't know what he's doing, that he is ϕ -ing, who he is, etc. – is someone who performed an action involuntarily because his ignorance was a decisive factor as to why he ϕ -ed (ibid.).

These factors Aristotle lists, by which someone can be said to be acting by reason of ignorance (1111a2-20), are redolent of the legal precedent established by the M'Naughten rules.³² This is a defence to exculpate the accused from guilt by pleading insanity. If the defence is successful, it blocks guilt *not* because the accused party has been justified (viz., has been found to have done the right thing or a permissible thing) nor because she has been found to have not done the act. Rather, if the defence successfully deflects guilt away from the accused, it does so because it manages to *excuse* her. This doesn't subtract from the fact that the accused has done something 'which is deplored' (Hart, 2008a, p.14). However, 'the psychological state of the agent when [s]he did it exemplified one or more of a variety of

³² A person is sane for purposes of ascribing legal responsibility viz., liable to public condemnation and punishment if: '1) he knows what he is doing and (2) he knows that what he is doing is, as the case may be, right or wrong' (Wolf, 2010, p.55).

conditions which are held to rule out the public condemnation and punishment of individuals' (ibid.). This is an excuse.

In cases where the M'Naughten rules apply, the defendant is excused because she is held to have acted involuntarily (or not fully voluntary) in virtue of her particular kind of lack of knowledge. Aristotle also seems to be speaking of insane actors when he writes 'of all of these no one could be ignorant unless he were mad' (1111a7). But clearly, it isn't only 'mad' people or people to whom M'Naughten rules are properly applied who can act involuntarily by reason of ignorance. Aristotle mentions Merope (1111a10-15) who killed her son believing him to be an enemy of hers. The crucial element in acting by reason of ignorant is that any action so done would *not* have been done *but for* the ignorance. And this is thought to *excuse* the actor thereby exculpating her from blame.³³ It is not the state of being ignorant as such that matters. It is whether that state is something that explains the performance of an act/action. So we must ask whether S ϕ -ed in ignorance or by reason of her ignorance.

For a similar reason, it makes sense to frame the discussion about excusing ignorance in terms of *blame* and *guilt* rather than responsibility. Austin states 'that it is to evade responsibility, or full responsibility, that we most often make excuses...' (1979, p.189). But we have just said, to excuse S, is *not* to deny something deplorable has been done and S contributed to its coming about. However, there is an ambiguity in responsibility such that to deny 'S is responsible for ϕ ' can be to deny that ϕ and S are causally linked at all. That is, to deny S is causally responsible for ϕ .³⁴ Whereas what we really mean when S is excused, is that there is reason to exculpate her from blame or guilt, not that she didn't have some part in producing ϕ . Likewise, 'S is responsible for ϕ ' can mean that S is *answerable* for ϕ (Smith, 2015, p.103). S is in a position such that we can legitimately ask S for a rational explanation for ϕ . This is consistent with having an excuse. It might be that in hearing S's explanation for ϕ , we discover an excusing condition e.g., S ϕ -ed by reason of ignorance (Baron, 2017, p.59).

³³ I'm not sure if we would excuse Merope for her killing. To me, it seems unlikely that she can be excused *tout court* either morally or legally. But there may be something in the neighbourhood for which she can be properly excused: given the killing was done by reason of ignorance, she is excused from the act of *killing her son* because but for her ignorance, she would not have done it. Whilst she may be excused from that, she *cannot* be excused from the act of unlawfully killing someone. (Unless we want to say Merope killed in pre-emptive self-defence which, if she were right about the identity of her victim, might *justify* her killing; not excuse it. But since she wasn't right, her act isn't justified, and neither can it be totally excused).

³⁴ See Hart's *Postscript* (2008a, pp.214f).

Finally, ‘S is responsible for ϕ ’ can also mean that S occupies a role such that S must *see to* ϕ (ibid., n.30).³⁵ Therefore, ϕ -ing by reason of ignorance excuses S by way of exculpating her from guilt, blame, or full blame.

This results in the following formulation of EI, EI’.

EI’: if S does ϕ but she does not know something which, *but for* that lack of knowledge, she would *not* have done ϕ , she ϕ -s by reason of ignorance which makes her ϕ -ing *involuntary*. Thus, S is exculpated from blame for ϕ -ing by way of excuse.

Yet EI’ is still a faulty account of exculpatory ignorance. Because even if S ϕ -ed by reason of ignorance, *and* we recognise S as potentially being excused from guilt or blame for ϕ , not for S’s responsibility for ϕ , we still need to distinguish between *culpable* ignorance and *non-culpable* ignorance. Presumably, only non-culpable ignorance can produce involuntary results which exculpate the agent involved in producing the results. Otherwise, one could be held not culpable for something that came about through one’s ignorance even though that person is culpable for being ignorant in just that way. That seems like an absurd consequence. For instance, Reuven feeds tahini cake to Shimon being ignorant of the fact that Shimon is seriously allergic to sesame. Whether Reuven is exculpated from blame for Shimon’s ensuing allergic reaction, depends upon whether Reuven is blameworthy or culpable for being ignorant of Shimon’s allergy. In turn, the culpability of someone’s ignorance depends on, *inter alia*, what sort of social role the ignorant party plays in the case. If they are strangers at a tea party, Reuven is likely *not* culpable for his ignorance and thus, served Shimon the cake by reason of non-culpable ignorance. This exculpates Reuven from blame. If, however, Reuven is Shimon’s primary school teacher, then Reuven may well be culpable for his ignorance.³⁶ After all, one of the demands of that social role is to be *in loco parentis*. Reuven hasn’t satisfied that demand if he doesn’t *know* the various foods which endanger his pupils. Therefore, we must add a non-culpability clause to exculpatory ignorance.

EI’’: if S does ϕ but she does not know something which, *but for* that lack of knowledge, she would *not* have done ϕ , and she is not culpable for that lack of knowledge, she ϕ -s by reason

³⁵ Ditto, (pp.212-214).

³⁶ The case is Baron’s (pp.58f).

of ignorance which makes her ϕ -ing *involuntary*. Thus, S is exculpated from blame for ϕ -ing by way of excuse.

Still, however, EI'' is false. If the antecedent is satisfied, the consequence – that ϕ is involuntary and thus S is excused from blame – does not follow. Unto here, we have presumed along with the standard view, that being ignorant that p is *only* inconsistent with knowing that p . Let us now suppose that Levi doesn't know about Shimon's allergy, but he does truly believe that Shimon is allergic to sesame. Levi ϕ -s (where to ϕ is to serve Shimon tahini cake).

- 1) The standard view of ignorance.
- 2) Ignorance exculpates (because it is a condition of involuntariness).
- 3) Levi ϕ -s.
- 4) Therefore, Levi is exculpated from blame for his ϕ -ing.

But Levi had a true belief about Shimon's allergy at the time he ϕ -ed! To truly believe that Shimon is allergic to sesame and yet serve him sesame regardless is, at best, grossly negligent and at worse, malevolent. So (4) must be false. As a result, the standard view of ignorance loses plausibility so long as (2) is true. The standard view produces the wrong verdict about doxastic attitudes which do *not* exculpate agents because of their ignorance. Other than (mere) true beliefs that p , attitudes such as non-Gettier-proofed but justified true beliefs that p , being aware that p , or a strong conviction that p are sufficient for being ignorant that p . But having a justified true belief that p which fails to be knowledge because it is prone to Gettier counterexamples, is not the sort of thing which makes the cognitive subject ignorant of p in the sense of her not being fully blameworthy for ϕ -ing (if she ϕ -ed because of her belief that p). "Levi truly and justifiably believed that Shimon was allergic to sesame, and so he served Shimon sesame" makes Levi appear *more* culpable, not less.

The *new view* takes ignorance to be the lack of true belief that p . This far better accounts for the well-established fact of ignorance's exculpatory power in laws and morals (Le Morvan & Peels, p.25). Therefore, I see no reason to think of ignorance as an exact complimentary to knowledge. Ignorance is inconsistent with knowledge, but it is also inconsistent with truly believing p and awareness that p (Baron, p.59). Exculpatory ignorance then, is a way of being

unaware of a proposition, but this unawareness is narrower in scope than thought by the standard view. It is *not* just a matter of not knowing some proposition. The ignorance that exculpates agents is a doxastic unawareness whereby the subject *lacks* the true belief that *p*.

EI'': if S does ϕ but she lacks true belief about something which, *but for* that lack of true belief, she would *not* have done ϕ , and she is not culpable for that lack of true belief, she ϕ -s by reason of ignorance which makes her ϕ -ing *involuntary*. Thus, S is exculpated from blame for ϕ -ing by way of excuse.

2.2.1. *Error and ignorance*

What is the relationship between acts done in error and acts done by reason of ignorance? Acts performed in error nearly always involve a component of ignorance. But they often do not acquire the excusatory force of EI. What are errors, though? They are not accidents. Accidents are events in which something happens (usually something unwelcome), but the actor involved in the event had no design whatsoever to bring about the occurrence (Daube, 2008, p.53). If a woodchopper's axe, whilst he is engaged in his chopping, flies from its handle and kills somebody, this is an accident (ibid.; Deut. 19.v). In accidents, the role played by humans are penumbral. There is no need for a human's rational powers nor their desires to get an accident to happen. There is no principled difference between the woodchopper's axe flying from its handle causing someone's death and a death being caused by lightning striking a tree and causing someone's death. Both are equally unwelcome and equally unlucky. Of course, it doesn't follow that in accidents involving human actors there won't be reactive attitudes which make it difficult for us to appreciate the parity between all types of accidents. We might find ourselves blaming D for V's death, even if this attitude is not appropriately held. Similarly, D himself might feel pangs of guilt because after all, he was involved. But D's feeling of guilt also does not necessarily correspond to the *fact* of his guilt. Accidents are horrible. But human agents' contribution to them is incidental for their powers are immaterial to the occurrence of the accident. Horrible they may be, but disregarding the possibility of negligence, accidents are 'misfortune[s], and nothing else' (Daube, p.64).

Errors, on the other hand, are made distinctive by the involvement of human's rational and conative powers. This is what makes them so tragic. In the limited sphere of human conduct

where we can hope to exercise freedom and harness the world to our designs, errors show us to be helpless even when everything goes to plan (ibid., p.65). Errors, then, are distinguished from accidents because they are marked by the human mind. An error is when a human agent carries out her plan successfully but there was faulty assumption in the plan. The ‘result is wanted, yet there is some defect in the undertaking as a whole, and the defect lies in a human mind, in the calculations on the ground of which action is taken (ibid., pp.64f). Therefore, when the Amalekite lad slew Saul believing David wanted this and would thereby spare him, he committed an error (2 Sam. 1; ibid., p.53).

The Amalekite’s error doesn’t excuse him. This is because of the legal doctrine that *ignorantia juris non excusat* (Black, 1999, p.749). We can speculate that the Amalekite either didn’t know the law that it was forbidden to kill God’s anointed (Sam. ibid., 1.14), or that he thought David would make an exception given his rivalry with Saul. Alas, for the Amalekite, this ignorance failed to excuse him. *Ignorantia juris non excusat* ergo David demanded his execution (ibid., 15).³⁷

Moreover, for an error to be exculpatory, the ignorance of the fact must be (i) earnestly held and it must be (ii) reasonably believed. For instance, Levi’s error in serving sesame to Shimon can only have exculpatory force if Levi (i) earnestly believed that Levi was not allergic to sesame and (ii) the belief is not unreasonable – and it is not unreasonable to believe a stranger is not allergic to sesame. This is made particularly stark in the case of *DPP vs. Morgan* [1976]. After a night of heavy drinking, accused brought home his friends informing them that all three of them could have sex with his wife while he, Morgan, watched. Accused also told his friends to not worry if his wife resists or says she doesn’t want to have sex. If she did that, she is only doing so to increase her pleasure. At the time, the House of Lords upheld as statute that a ‘man commits rape if –

(a) he has unlawful sexual intercourse with a woman who at the time of the intercourse does not consent to it; and

³⁷ Does ignorance of the moral law excuse? This is a complicated topic, and one I cannot answer in full here. For what it’s worth, I am split. On the one hand, ignorance of moral law matters because an agent’s quality of will matters, and to be the sort of person who lets himself be ignorant of the moral law is objectionable. This is a kind of ‘blindness’ or failure to grasp the ‘meaning of life and man’s position’ which *inculcates* agents (Daube, p.74). On the other hand, I am sympathetic to the view that there is something unfair about blaming someone whose non-culpable ignorance of the moral law, involving a belief genuinely held that ϕ -ing is morally permissible (where ϕ is some immoral act), led to her ϕ -ing. See Zimmerman (2002) who argues that ignorance of wrongdoing always exculpates wrongdoers from blame.

(b) at that time he knows that she does not consent to the intercourse or he is reckless as to whether she consents to it.’³⁸

Because (b) stipulates the accused *knows* that the woman did not consent at the time of intercourse, it comes to preclude someone who *believed* the woman consented, if that belief was held earnestly and reasonably. Pursuant to the judge’s direction to the jury, that they must consider whether the defendants honestly believed that Mrs. Morgan consented and if they did, whether this belief was reasonable (Duff, *ibid.*), the defendants were not acquitted. Thus, one can plausibly infer that the jury found against the defendant’s claims that they were honestly and reasonably ignorant of Mrs. Morgan’s not consenting. This case illustrates that at least in 1976, being ignorant ‘earnestly’ and ‘reasonably’ sufficed to exculpate someone from the crime of rape. My mentioning of *DPP vs. Morgan* is not meant to be interpreted as me endorsing the Upper House’s ruling, obviously. It is, however, a vivid exhibition of how error can be seen to operate with exculpatory force in the law. (Having ignorance of this kind – (i) and (ii) – functions as a legal excuse with crimes other than rape).³⁹

Finally, to exculpate an error because of its component of exculpating ignorance abolishes or mitigates an agent’s guilt or blameworthiness. But this is *not* to say that agents who perform errors worthy of exculpation have nothing to rectify. Leviticus proscribes many actions, not actions as such, but prohibitions are enforced on actions done by a person in a particular condition or on objects of action belonging to special categories (Daube, p.72). For example, entering the sanctuary as such is permissible; entering in a condition of impurity is prohibited. And when one commits an error e.g., they unwittingly enter the sanctuary in an impure state, and that error is made known to them, they owe rectification (e.g., Lev. 4). In this case, a sin offering must be brought. This is congruent with the claim that errors done because of ignorance may save you from full-blown guilt or blame, but not from ‘pollution’.⁴⁰ The idea being that although the law (or our moral reactive attitudes) recognises ‘in ignorance’

³⁸ Quoted in Duff (1981, p.49).

³⁹ See Duff (*ibid.*, especially pp.55f; 58f) for why rape is a categorically different crime to assault and why this means conditions of negligence don’t apply in rape. Instead, a man who went ahead with having sex on the reasonable belief that his partner *might* consent to it but also *might not* consent to it, should be considered *reckless*, not negligent.

⁴⁰ Nahmanides (Lev. 4.2, s.v. *nefesh ki teheta vishgaga*) employs this model of understanding sins done unwittingly: these sins ‘pollute the soul and are a blemish on it’. See Lewinsohn (2006, pp.100f) for discussion. Also, see Hyman (2016, p.694) who responds to Sharon’s assertion that though ‘Oedipus was ignorant of his familial relationship to Jocasta, and yet he bore the guilt of mating with his mother’ (p.681) with the charge that Sharon confuses guilt and pollution. Oedipus bore no guilt, but he may have been sullied.

extenuating circumstances in such a way that it blocks an inference from bad act to bad agent,⁴¹ it cannot *undo* what was done. There is still some ‘guilty’ or ‘blameworthy’ remnant which stains agents who acted in error. It is coherent for legal systems and persons to simultaneously excuse an agent on account of ignorance whilst also expecting the agent carries out an act of expiation for her infraction. If I am rational, I won’t resent the person who failed to turn-up to the meeting because she earnestly and reasonably believed it was cancelled. But do I not deserve an apology? Needing to make amends is perfectly compatible with exculpating someone by way of excuse. Acting by reason of non-culpable ignorance makes acts involuntary which subtracts guilt and/or blame, but not necessarily the actor’s need to expiate herself.

2.3. *The compulsion condition*

Compulsion is a kind of force. It is a force that makes doing something non-optional, that is, compulsory. So, when a subject is compelled to ϕ they have no option but to ϕ . But there are many cases where something has no option but to ϕ , where it would be a violation of our understanding of the concept of ‘compulsion’ to describe these things as being compelled. For example, the beating of my heart is non-optional, but it seems wrong to say that my heart is compelled to beat.⁴² However, if compulsory means no more than non-optional, there *is* a sense in which it is true to say my heart is compelled to beat. My heart is compelled to beat, the pen is compelled to drop when I let it out my grasp, and the billiard ball is compelled to move when it is struck with sufficient force. Let us then distinguish between ‘ x is compelled to ϕ ’ and ‘ x is under compulsion to ϕ ’. Being *under compulsion* conveys a more accurate sense of the compulsion condition which opposes voluntariness. This is because it gives the impression

⁴¹ A paraphrase of Rosen’s (2008) characterisation of excuses. An excuse is ‘any consideration which blocks the normal inference from bad act to culpable agent’ (p.592).

⁴² It might be objected that ‘non-optional’ is a misnomer applied to things like heart beats and billiard balls. Coming from the Latin ‘optare’ to choose, ‘non-optional’ is an agential notion so is only appropriate when there are options. (I thank O’Brien for this query). This is reminiscent of Wittgenstein’s suggestion that the tokens of involuntary movements are those things which don’t obey orders as voluntary movement does (1967, §593). My beating heart cannot obey an order to beat – it simply just beats. However, the objection succeeds only if we are prepared to accept a view of agency which I don’t accept (and face no pressure to accept). Namely, that only choosing, rationally responsive human beings are agents. Another paper is required to properly vindicate this thought, but I am quite happy to apply agency to beating hearts – they cause change by pumping blood around my body. (See Hyman (pp.29-32)). Granted, beating hearts don’t have conscious options. They don’t execute decisions. Yet beating hearts also eventually stop. So, they both beat and not beat. They also skip beats and have attacks. These are, in a sense, options the heart has even though hearts don’t execute their decisions to choose like human agents do. Therefore, my heart’s *beating* is non-optional because it cannot follow orders to beat – it just does. But hearts have ‘options’ because beating isn’t the only thing they do; they eventually stop beating. Consequently, ‘beating hearts are compelled to beat for it is non-optional’ is felicitously rendered.

of a subject being *subdued* or *yoked* with an exterior, oppressive force which the subject under compulsion doesn't welcome. Thus, whilst many things and processes are compelled to do what they do, not everything which is compelled to do what they do is under compulsion to do what they do. Hearts are compelled to beat whereas, arguably, the drug addict is under compulsion to take drugs.

The distinction is relevant because it assists us in latching on to an important feature of compulsion. Namely, that compulsion *depends on the desires, preferences, and rational powers of agents*. The distinction between my beating heart and the drug addict isn't the difference between a non-agent and agent. Rather, it is the difference between what sort of psychological constituents (or none) typically make-up the agent in question. Beating hearts, billiard balls, and vending machines can all be agents. But their agency isn't coextensive with desires, preferences, and rational powers. The vending machine cannot desire to spew out all its snacks, nor can it prefer to not give me my KitKat when I push the right combination of buttons, and nor can it think about the best means to achieve its goal of handing me my KitKat. And if there is no gap or tension between an agent's desires, preferences, or rational aims, an agent *cannot* be under compulsion. The most they can be is compelled.

Competent human agents are not agents because of humans' psychological faculties. But typically, agency in human beings is *realised* through those psychological faculties. By this I mean that usually, when humans act in a way that is attributable to them as agents, desires, preferences, rational planning and the like are bound to feature. That is, when we act, we frequently do so for *reasons*. In competent human agents, agency operates through the prism of these psychological faculties; faculties which seek or determine our reasons for acting. When I cause the chair to move I don't just cause the chair to move like the billiard ball's collision with another ball causes the latter to move. I've moved the chair because, say, I *wanted* the chair moved or I *wanted* to move the chair. It is the function of these sort of faculties which allows for the possibility of being *under* compulsion. This occurs when there is a gap between the agent's wants, desires, preferences, and the compelling force. This results in the following formulation which captures the root of the matter.

To be under compulsion (UC)⁴³: to be in a condition such that there is tension between what you want and what must be done.

‘Want’ is to be interpreted widely, standing for all and any of the desires and preferences the agent has at the time of compulsion which is in tension with the demands of the compelling force.⁴⁴ And the ‘what must be done’ is the demand issued by the compelling force. It is not ‘any old’ ‘what must be done’.

So, it is a necessary condition of UC that the agent subject to it has wants. Moreover, and crucially, these wants are the product of rational powers – they are the sort of thing which can be justified by giving a reason. This means that on my account of UC, creatures which have wants but no reasons for those wants cannot be UC.⁴⁵ It is not obvious that Frankfurt’s wanton, for example, can UC (1998b, p.17). This isn’t because the wanton lacks rational powers – she can deliberate about how best to satisfy her wants – but because she is indifferent regarding the reasons for her desires. Presumably then, she will be inclined to desire whatever thing presents itself most intensely *qua* desire. Since courses of action issued under compulsion present themselves intensely – they are proposals of action which dominate the psychological horizons of an agent when she considers her possibilities because they are things which must be done – it is likely that the wanton will always end up desiring the proposal issued by the compelling force. In which case, the wanton’s preclusion from UC is a contingent fact, rather than a necessary feature of their psychological make-up. Possessing no reasons for her desires, however, the wanton might be compelled to pursue whatever desire she happens to have in much the same way the operations of a vending machine are compelled.

⁴³ ‘UC’ means the state of ‘being under compulsion’ and should be read as so, modifying or deleting the existential verb ‘to be’ as necessary.

⁴⁴ Sometimes in this section it is also used interchangeably with ‘desire/s’.

⁴⁵ Don’t object that this account of UC is implausible because it rules out animals from UC. I agree, I think that would be an objection if true. But it isn’t true. Even though Jamjam (my dog) cannot himself supply reasons for his behaviour and wants, a little attention to his actions show they are reason-based. His actions reveal their means-end form. For instance, he isn’t pattering on the glass window just so. He does that *in order* to get our attention *in order* to get the door opened. So Jamjam *can* UC. (Although, to qualify, Kenny may be correct when he says animals like Jamjam aren’t conscious of their ends *as ends* (1992, p.39)). I admit that my imputation of rational powers as a necessary condition of UC doesn’t sit very nicely with the fact that non-human animals are also candidates for UC. But it is a cosmetic impairment rather than one of substance. To mollify the cosmetic damage a longer dissertation on rational powers would be required. I can’t do that here.

Given UC's dependence on desires, preferences, and rational powers, it is inherently an evaluative notion. The same show of force may for some people precipitate an act done under compulsion whilst for others it may be no such thing. Thus, whether ϕ was done under compulsion depends on the subject's desires and on our (i.e., onlookers') assessment of the subject's desires. To illustrate, consider the inverse-bouncers. In this case, the bouncers let anyone enter the club but exiting before the weekend is over is strictly forbidden and the bouncers enforce it. They will use overwhelming and irresistible physical strength if needed to keep everyone inside the club until Monday morning. Mark and Jez are both inside the club. Mark wants to leave. He even *tries* to leave. But he can't. Clearly, he is under compulsion to stay. Jez, however, is the partying sort. He can't think of a better way to spend his weekend than to be stuck in a club for its entire duration. Jez is so engrossed with his dancing, he hasn't even noticed the inverse-bouncers. So, although Jez's circumstances are the same as Mark's, Jez clearly is not UC to stay whereas Mark is UC to stay. The reason is because Mark wants to leave whereas Jez doesn't want to leave – the thought hasn't even occurred to him. Alternatively, we can say Jez is compelled to stay – he after all cannot do anything but stay – whereas Mark is UC to stay.

The above is of course just a *Peep Show* corruption of Locke's man in a locked room case.

Again, suppose a man be carried whilst fast asleep into a room where is a person he longs to see and speak with, and be there locked fast in, beyond his power to get out; he awakes and is glad to find himself in so desirable company, which he stays willingly in, i.e. prefers his stay to going away. I ask, is not this stay voluntary? (*Essay*, 2.21.10)

I agree. The man's stay is voluntary because he wants to remain in the room. Thus, Locke resoundingly and correctly concludes that '*voluntary* then is *not opposed to necessary, but to involuntary*' (ibid., 11). Denying that voluntariness opposes necessity is another way of stating that doing what must be done is insufficient for UC conditions to apply and thus, insufficient to make an act or omission involuntary. A conjunction of a compelling force plus a conflict between the demands of the compelling force and the wants of the agent is required for the compulsion condition.

Finally, compelling forces are of two kinds. (1) Forces which *obstruct* a person's will and compel a person by militating *against* their will. (2) Forces which *hijack*⁴⁶ a person's will and compel a person by working *through* their will. Coercive threats from other (human) agents typify (1), where the aim is to get the person UC to do something they don't want to do. Furthermore, physical environmental factors can also constitute UC (1). In these instances, it would be wrong to speak of the compelling force's *aims*, since environmental factors like weather don't have aims, though exposure to the elements can be compelling forces. For example, having to trespass to protect oneself from a storm (Hyman, p.81) or having to swim because one slipped and fell into a deep pond (Wittgenstein, 2009, §613).

Psychical urges which make a subject do something because on some level (first-order or higher-order) or under some description they want to do it, exemplify type (2) compulsions. Instances of (2) which count as exculpatory compulsions are the unwilling drug addict (she wants to take drugs, but she doesn't want to want to take drugs) and people with compulsive disorders e.g., the compulsive handwashing of someone with OCD. The main difference between (1) and (2) is that in (2), the *motive* for the action comes from within the agent herself whereas (1) instigates a *motive* for action from the outside.

However, this distinction is not clear-cut, and they are not mutually exclusive. They can even be seen as two sides of the same coin for every coercive threat does two things: it (i) threatens and (ii) offers to protect or advance an interest of the victim. (i) is the sanction that I will incur if I violate the demand of the compelling force. (ii) is what I will get *in return* for satisfying the demand. The presence of (i) means coercive threats are type (1) since an outside source gives me motive to do something I do not want. But the presence of (ii) means the same coercive threat is also type (2) since I now have the motive to do the act on account of the protection of my interest. E.g., When Yehuda blackmails Issachar, the threat – like virtually all threats – is biconditional: “I shall release the career-crushing information I have on you iff you don't A.” The content of (i) is the releasing of the sensitive information. It externally motivates Issachar to A. The content (ii) is not releasing the information which amounts to the preservation of Issachar's reputation. This internally motivates Issachar to A.⁴⁷

⁴⁶ If memory serves me correctly, I owe credit to Carla Bagnoli for the selection of this most apt word to describe compulsions of type (2). By *viva voce*, April 2022.

⁴⁷ See Nozick (1997, pp.23-31) and Frankfurt (1998a, pp.28ff) for discussion delineating the difference between offers and threats.

2.3.1. *Compulsion, choice, and evaluation*

Let us now refocus on Locke's denial of necessity's opposition with voluntariness because there is another important lesson in this. For the compulsion condition to apply, a case does *not* need to feature necessity, if necessity means something like the following. If B necessarily follows A then whenever A, we have B; B cannot fail to happen if A happens. Call this *strict necessity* (SN). One might think that because UC is understood partially in terms of 'what must be done', bona fide cases of UC leave no room for choice. If the compelling force demands A, B invariably happens. But, in fact, UC doesn't require this mode of necessity, SN. Note how this entails that compulsion doesn't abolish choice. Since B isn't strictly necessary, the force presented by A doesn't supply the agent subject to UC with the choice between $B \vee \neg B$, that is, no choice at all. At first blush, this may be difficult to accept because voluntariness *is* in an important respect about choice. How then does UC get to cancel the voluntariness of agent's doings/omissions if it doesn't also cancel choice?

There is a distinction between *literal* choice and *real* choice. A literal choice is when there is at least more than one possible course of action open to the agent. A real choice is when not only is there more than one possible course of action open to the agent but taking either of these courses doesn't bear too high a price. Hence, when someone merely has literal choice to A we say, "she has no real choice to A". And if we say in these circumstances that "she has no choice" what we mean is that she has no real choice (Hart & Honoré, 1985, p.41). Cases of duress generally take the form of supplying a subject with literal choices but not real choices. For instance, when the mafia get Zebulun to commit a crime under threat of life and limb, Zebulun might be able successfully to plead duress thereby being exculpated from guilt. So, if duress is considered a species of necessity, we must be cautious to avoid considering it a species of SN. Rather, duress is a species of 'hypothetical' necessity (HN) (Hyman, p.82). Duress is a form of HN because it presents agents with necessary courses of action *if* some other end is to be avoided or achieved. If I am to make it to work on time, then it is hypothetically necessary that I wake up at 7am. It doesn't follow that I have no choice in the matter. I can quite easily choose to wake up late (or fail to wake at 7). Episodes of duress supply the person subject to it with strong influences on her will to eschew personal harm or the harming of someone closely associated with her (e.g., family members).

Yet, it manages to do this without absolutely eliminating choice. This is because Zebulun had the *literal* choice to either submit to the threat or resist the threat and suffer the consequences. Law makes room for cases of duress because it recognises that for most reasonable people under ordinary conditions, suffering the consequences by paying with one's life and limb is too high a price to pay. Usually, people are not expected to undergo that suffering to eschew criminal activity. In the words of Lord Morris, ““The law must take a commonsense view and remember that the instinct and perhaps duty of self-preservation is powerful and natural”” (quoted in Kenny, 1982, p.199). But the law is limited in its accommodation of this ‘powerful and natural’ instinct, for it cannot tolerate the instinct if the induced act is so serious as to outweigh the drive for self-preservation. Treason, for instance, is likely a crime too hefty to allow for pleas of duress (*ibid.*, p.198). Therefore, it is not only the presence of coercive threats which matters to voluntariness. Our ascriptions of ‘in/voluntary’ to acts done also depend on the *nature* of the act the victim of the threat is induced to perform.⁴⁸

This principle is transferable to moral conduct, where a subject is under duress to violate some moral rule (not just or not only a criminal law). Here too the *nature* or the *particulars* of the immoral act subjects are induced to perform are material in determining whether subjects, if they give way to the threat, act in/voluntarily. For example, Aristotle insists it is absurd to think that Euripides’ Alcmaeon was compelled to kill his mother (*NE* 1110a25-30). Brown remarks that Aristotle is probably asserting that Alcmaeon cannot justifiably claim UC to modify our moral assessment of his act (2009, p.218). Brown’s commentary doesn’t go far enough. It isn’t only that there is no justice to Alcmaeon’s claim of being UC. It is that *ab initio*, there was *no compelling force*.

The underlying thought is that no reasonable person can be compelled to slaughter their own mother, no matter how grievous or violent the threat. That is to say, Alcmaeon is presented with a *real* choice between surrendering to the threat or suffering the consequences of resisting it. The price – and let us assume it was Alcmaeon’s life – is *not* deemed too high, too

⁴⁸ There are sceptics or revisionists of legal duress. This position reaches its logical climax in Sir James Stephen’s contention that there is no crime borne from duress that the law can legitimately excuse since the ‘criminal law is itself a system of compulsion on the widest scale’ (quoted in Kenny, *ibid.*, p.204). The law, serving as a body of threats to confiscate property, repeal liberty, and take life, cannot withdraw its threat the moment opposing *per minas* threats appear.

For wider discussion, especially about whether duress ought to be a complete defence or a mitigation (or neither), see Kenny (*ibid.*, pp.197ff). For Kenny’s views on the law’s power to deter criminality done under duress, see (*ibid.*, pp.204f).

burdensome, compared with the act he is induced to perform – murdering his mother. This is what Aristotle means when he says, ‘some acts, perhaps, we cannot be compelled to do’ (*NE* *ibid.*). It is *not* that there is no coercive element in these cases. It is that the coercive ingredient is insufficient to count as a UC compelling force. This is because ordinary agents are morally (or legally when law is relevant) expected to not be so influenced by the coercion that they go ahead with doing the act the tormentor wants done. Coercive threats then, are neither sufficient nor necessary for compulsion conditions to apply. They are insufficient because like with Alcmaeon, someone can face coercion to A without that meaning he is UC to A. They are unnecessary because it is possible to UC purely by way of (2), when it the subject’s own will and motivations – not any inducement from foreign human agents – making her A. (Here, it can only be said metaphorically that the subject is coerced). Hence, UC is consistent with literal choice but not with real choice, and some coercive threats, despite their coerciveness, present agents with real choices.

Thus, we can conclude our exploration of the compulsion condition by emphasising how evaluative notions and principles pervade the concept of voluntariness.

- (a) It is evaluative because voluntariness depends on agents’ choices which are to do with the sort of things they judge to be preferable or desirable.
- (b) It is evaluative because our employment of the concept in law depends on juridical policy decisions regarding what coercive threats or compelling forces reasonable agents are thought to be able to withstand, and which ones they cannot.
- (c) Also, voluntariness is evaluative because it depends on our moral assessment of the case. Whether it is morally expected for reasonable agents to yield to a coercive threat/compelling force, or whether resisting is morally supererogatory or heroic.

An ethical concept by design, rather than a tool to aid us pin-pointing the causes or provenance of action as opposed to mere movement, it is no surprise that evaluative features and norms are in-built to voluntariness. Fashioned to respond to problems arising in the ethical dimension of human action, we should expect nothing less.⁴⁹

⁴⁹ There are also distinctive *subjective* and *modal* features to voluntariness which differentiate it from abilities and possibilities. I lack the space to explain and disagree with some of Hyman’s points on these issues. But see Hyman (2015, pp.93-9) for discussion. See Alvarez’s (2016, p.671) second objection to Hyman’s account of the relationship between voluntariness and obligation. It is an animadversion simpler and more effective than her first.

2.4. *Voluntariness and the will*

The claim that voluntary acts are to be identified with movements resulting from acts of will, volition, or conations has been dispatched in the first section of this Part. But often when we talk of ‘the will’ we are *not* talking of discrete *acts* of will or *willings*, but something else entirely. Indeed, I have already used it in this sense in the preceding section. What we (at least philosophers) often mean by ‘the will’ is the aggregation of salient aspects of a person’s psychology including (but not limited to) their beliefs, desires, cares, concerns, commitments, and values. Decisions and actions flowing from the will is privileged *not* because they are the result of some endogenous mental cause in the agent, but because there is some special relation of *belonging*, of being *owned* by the person whose will it is (Hieronymi, 2022, pp.14f). I shall follow Hieronymi by presenting this quote from Frankfurt: ‘...what he cares about is far more germane to the character of his will than the decisions or choices he makes’ (1998c, p.84).

The will so conceived is impervious to fluctuations between voluntariness and involuntariness. I don’t mean that the will isn’t affected *tout court*. But even in (1) UC cases where subjects’ wills are obstructed, wills remain intact. Being obstructed or deflected doesn’t entail the will’s destruction or being overborne. In fact, some genuine cases of compulsion may give *fuller* expression to the will’s integrity than it would otherwise be shown. Consider a victim UC, who surrenders to coercion to spare her family from harm by doing A (the act she is induced to perform). It is true she has no desire to A, but she is resolutely committed to her family’s welfare. Thus, by *doing* A viz., by doing what, by many people’s calculation, is foremost a subduing of her will, she demonstrates a healthy and potent will. The will *isn’t* eliminated by involuntary acts. By the same token, not all voluntary acts reflect persons’ wills because not every voluntary act we do can give expression to our will. My intentional finger-tapping is surely an act attributable to me and to my desire, but it doesn’t represent any of my cares or values. It is just irrelevant to my will. ‘And the will does not produce sneezing, *nor yet walking*’ (Wittgenstein, 1967, §579).⁵⁰ Dispensing with the erroneous thought that willings are mental causes of action or ‘direct, non-causal producing’s of action (ibid., §580), evinces that semi-autonomous actions (sneezing) and actions like walking or finger-tapping are alike in this way:

⁵⁰ My emphasis.

neither of them carry the mark of the will. Drinking water is quite unlike my writing a philosophy thesis in this sense; my will is channelled through this act.

Aside from this, vernacular uses a derivate of the proper noun ‘the will’ to describe an attitudinal orientation towards acts. We say an act was done either ‘willingly’ or ‘unwillingly’. Neither of these attitudes correspond to the distinction between in/voluntary acts (Hyman, p.84). ϕ -ing willingly contrasts with ϕ -ing reluctantly (ibid.). And I can willingly do something *involuntary*, as when a mother willingly surrenders to a coercive threat to spare her family from harm. And I can *unwillingly* do something voluntary, as when I abstain from food to lose weight. Involuntariness neither shows the absence or subjugation of the will nor does it show that an act was done unwillingly.⁵¹

2.5. *Voluntariness and control*

Many assume that voluntariness involves a control condition. Where there is no control over ϕ -ing, ϕ wasn’t done voluntarily. And because voluntariness has been mistakenly identified with agents’ *acts*, control is also often thought to be a necessary condition of responsibility.⁵² But we don’t need reams of citations to give these claims support. They already carry intuitive plausibility if we accept some form of the view that makes voluntariness a condition of agency. The thought goes: I can only act if my movement is voluntary, and it is only voluntary if I control my movement. If I am responsible for my acts, control is necessary because part of what it is to act is to exercise *voluntary control* over my movement. This phrase, ‘voluntary control’, is ubiquitous. I don’t deny that there is anything useful to this refrain, that it doesn’t latch onto something philosophically significant. Only, I confess that ‘voluntary control’ is, to me, obscure. Consequently, I am unclear about the precise relation between voluntariness and control. To users of the phrase, the relation seems apparent. Namely, the relation appears to be *biconditional*, such that whenever we find instances of controlled action, we also find instances of voluntary action and *vice versa*.

⁵¹ *Pace* Aristotle (*NE* 1110b19, 1111a20) and Anscombe (2000, §49) who hold that feelings of reluctance or regret reveal an act’s involuntariness. Hyman rejects this completely (ibid.). I am convinced that acts done willingly says nothing about the act’s voluntariness. I am less convinced about Hyman’s claim regarding reluctance/regret, and involuntariness. Not because feelings of regret decisively show an act’s involuntariness. That is dead wrong. But feelings of reluctance/regret oftentimes go some way in intimating an act’s involuntariness because they can be affective consequences of *discovering* one has done something involuntarily, say, by reason of ignorance.

⁵² E.g., Wolf (1990, p.44).

It seems to me that control is a conjunction of powers. S has control over her ϕ -ing *just in case* S possesses the powers to

- 1) Initiate ϕ -ing.
- 2) Modulate ϕ -ing.
- 3) Inhibit ϕ -ing.⁵³

Thus, my view is that control is something which is frequently coextensive with voluntary acts. But voluntariness *doesn't* imply control, and control *doesn't* imply voluntariness. So I deny that they are logically equivalent. Now, there is probably a reason why voluntariness and control generally go together. That is, it is surely not a coincidence. Still, they can come apart.

- a) I can command powers (1-3) over my ϕ -ing whilst I ϕ *involuntarily*. Plainly, this occurs in most cases of involuntary action where agents ϕ by reason of exculpatory ignorance or because they are UC. The mother who ϕ -s in submitting to a coercive threat to save her family, is in complete bodily control of her movements when ϕ -ing. If ϕ -ing is robbing a bank, she can do all of (1-3) throughout the process of ϕ -ing. Her bodily movements which make her ϕ are fully up to her. Yet, she has still ϕ -ed *involuntarily*.
- b) A little harder to imagine, but not impossible, are cases where I act voluntarily though I *lack* (1-3). There are compulsive conditions where people have no control over some of their bodily movements because they cannot initiate, modulate, and inhibit them. Being a little fast and loose, let's consider Lady Macbeth's handwashing in Act 5.1 to be properly compulsive. Here I mean compulsive in such a way that vitiates control. Though she is somnambulist in this scene, we can imagine that her compulsive handwashing also takes place in her conscious hours. Even when she wakes it shall remain, in the words of the Gentlewoman, an 'accustomed action with her'. In these conscious instances Lady Macbeth lacks control over her actions, but it doesn't follow she handwashes *involuntarily* because doing something compulsively doesn't mean that the compulsion is the *motive* for her handwashing. Just like Jez is compelled to stay in the club, it doesn't follow that that is his reason for doing so. Perhaps Lady Macbeth really enjoys her accustomed handwashing that if she lacked the compulsion,

⁵³ I am indebted to Hyman who shared his thoughts on control with me (personal conversation on 24/03/2022), though they do not exactly align with my own.

she would try her very best to conjure up other means to achieve the same result. Therefore, the operative reason for her handwashing isn't her compulsion. Hence, she lacks control over her handwashing (because her compulsiveness undercuts control) and she handwashes voluntarily.⁵⁴

In sum, there is no biconditional relation between control and voluntariness. Control doesn't imply voluntariness (a), and voluntariness doesn't imply control (b).

2.6. *Voluntariness and intention*

Anscombe states that 'every intentional action is also voluntary' (§49). Williams makes a very similar recommendation in (1995, p.25) and (1993, p.66) where he writes voluntary acts are done 'if [they] are intentional aspects of action done in a normal state of mind'. But we should now be confident that intentional action need not be voluntary. Whenever I unwittingly serve sesame to a guest with a sesame allergy, I serve the sesame intentionally. (Although under some description I do not e.g., I don't intentionally serve her something she is allergic to). Moreover, whenever someone gives way to a coercive threat or duress, they do so intentionally. When a prisoner reveals state secrets under torture, they intentionally do so to stop the torture (Hyman, 2016, fn.18). When a mother commits a crime UC, she intentionally does so to save her family from harm. But neither of these cases involve agents acting voluntarily. Thus, the presence of intention doesn't guarantee voluntariness. Also, acting in involuntary conditions doesn't undercut the possibility of acting intentionally.

2.7. *Voluntary agency and voluntary passivity*

Thus far we have been considering voluntariness as an aspect of agency viz., voluntary *action*. A considerable virtue of Hyman's account of voluntariness is that it recognises that voluntariness is a description applicable to *passivity* as much as it is to activity. Agents can

⁵⁴ (b) is similar in structure to agents who willingly and voluntarily discharge their obligations (if it is true that obliging factors are usually compelling factors). (See Hart and Honoré (p.335) who state that carrying out intervening moral obligations doesn't negative the causal connection of the original harm inflicted on victim by the first actor. Ergo the wounds sustained by victim during third-party's discharge of obligation is legally discounted). Because what matters in determining whether someone acted (in)voluntarily whilst subject to obligation, is their *motive* for acting. If they acted *not* to discharge an obligation but for some other reason, that reason makes it the case that they fulfilled a compulsory action voluntarily. Loving parents generally do not look after their children because they have an obligation to, but because they love them. Even though they accept that they do in fact have an obligation to look after them (Hyman, p.96).

undergo actions or events without ignorance and/or compulsion, or with ignorance and/or compulsion. This possibility was in the main ignored due to the influence of volitionism/conationism (Hyman, pp.9f). Anscombe nods to voluntary passivity with her admission that things can be voluntary which are ‘not one’s doing at all...as when someone on the bank pushes a punt out into the river so that one is carried out and one is pleased’ (ibid.). Voluntary passivity’s kernel is a certain frame of mind⁵⁵ on the part of the patient. The patient must *consent* to the act or event being done unto them for it to qualify as an instance of voluntary passivity.⁵⁶

For example, if Pharaoh really wanted his heart hardened so that he consented to the hardening of his heart (Ex. 9.xii), Pharaoh underwent the hardening of his heart voluntarily.⁵⁷ But his heart *being* hardened by another actor is no act of Pharaoh’s. Likewise, I can voluntarily hear – providing hearing is no action and is a state of receiving audial sensory input – the music from my neighbour’s garden, if I am aware I can easily avoid it if I so chose. But I can also be subject to it involuntarily, where I have nowhere else to go and I know the music won’t stop. One can undergo surgery voluntarily, be kissed voluntarily, and fall asleep voluntarily (Hyman, pp.10f). In all of these, the person is a patient; the object of an action done unto them. Therefore, no reason exists to restrict our attributions of voluntariness to cases of activity.

2.8. *Voluntariness’ minimal condition*

EI” and UC are conditions for involuntariness. If voluntariness is ascribable to *anything* that is not EI” and UC, then it could be ascribed to almost anything. And then, it wouldn’t be intelligibly applied and the concept would lose its use. Thus, we want to limit the sort of things which can be voluntary. The conditions of *involuntariness* instruct us as to the basic qualifying condition of voluntariness. Because EI” is about *beliefs*, and UC is about *choice* and *desires*, to be a candidate for voluntariness and involuntariness, the subject needs to be a *doxastic* creature – someone that has beliefs about the world – and an *evaluative* creature – someone that has wants, preferences, and a capacity for choice. Only doxastic and evaluative

⁵⁵ I’m neutral as to whether consent is an attitude or an act.

⁵⁶ See Hyman (ibid., pp.88ff) for more on consent and voluntary passivity.

⁵⁷ See Stump (1988, p.419) for excellent Frankfurt-style discussion and interpretation.

creatures are intelligible candidates for doing (undergoing) voluntary and involuntary acts and events.

2.9. Conclusion

The challenge of distinguishing between mere movement and action generated an influential theory of action. Variations of this theory can be grouped together on account of their postulation of mental causes (or sources) in the form of conations, volitions, and acts of will to explain the boundary between mere movement and action. The deleterious consequence for voluntariness was that its true meaning and function was occluded by it being subsumed into this theory of action. Voluntariness was conflated with action itself. Such a conflation also compromised the individuality of different axes of human conduct. Action belongs to the axis of agency, which is to do with powers that bring about change. Voluntariness belongs to the ethical dimension of human conduct. Its use and design are to assist with legal and ethical assessment, especially with determining the innocence or guilt of an agent.

Following Hyman, I elaborated and defended the method to theorise about voluntariness via negation. To think backwards from what sort of things are *not* called voluntary and why, to apprehend the nature of voluntariness. Being a doubly-negated concept, voluntariness applies to cases where the agent (patient) is *neither* EI'' nor UC in their ϕ -ing (or in their being ϕ -ed upon). This is because EI'' and UC are typically factors which exculpate agents (patients) from wrongdoing. I proceeded to elucidate the evaluative nature of voluntariness – an evaluative nature for a concept designed to help us evaluate. I considered the repercussions that our account of voluntariness has on the will, control, intention, and passivity. Lastly, I explained the minimal condition for voluntariness which corresponds to the conditions for involuntariness. All these lessons will be crucial in understanding the dynamic between voluntariness and belief.

Part 3: Doxastic Voluntarism

I will argue for a strong iteration of doxastic voluntarism. Not only is doxastic involuntarism theoretically possible, not only is it sometimes true about a small selection of beliefs, but given the way we form and possess beliefs is neither done by reason of ignorance nor under compulsion, doxastic voluntarism is *routinely* true.

I open by formally stating the theses of doxastic voluntarism and involuntarism (3.1.). I then explore the doxastic ignorance and compulsion conditions for involuntarism (3.2.-3.3.4.). I argue that these conditions routinely *don't* impinge on our beliefs. I end by arguing that my account of doxastic voluntarism avoids the thorny issue of doxastic *agency*, and whether beliefs are *acts* or *states* (3.4.). This is a virtue of my theory: voluntarism can apply regardless of whether we perform doxastic acts or undergo doxastic states. Finally, I contend with a *prima facie* forceful objection. I respond by distinguishing between wants and wishes (3.4.-3.5.).

3.1. Doxastic (in)voluntarism and (in)voluntary belief

Doxastic voluntarism is a thesis about the relation between belief and voluntariness. It isn't about the relation between belief and the will, belief and intention, or belief and control. Now we know what voluntariness *is*, we can clearly grasp what doxastic voluntarism amounts to.

A reminder of the conditions for involuntariness.

EP': if S does ϕ but she lacks true belief about something which, *but for* that lack of true belief, she would *not* have done ϕ , and she is not culpable for that lack of true belief, she ϕ -s by reason of ignorance which makes her ϕ -ing *involuntary*. Thus, S is exculpated from blame for ϕ -ing by way of excuse.

UC: to be in a condition such that there is tension between what you want and what must be done.

Voluntariness = \neg (EI'') and \neg (UC).

Therefore,

Doxastic voluntarism (DV) = \neg (EI'') and \neg (UC) when epistemic agents (patients) believe that p or form their belief that p .⁵⁸

Consequently, doxastic voluntarism is true just in case it is possible for epistemic agents to believe p without those two conditions applying in their believing p . And for every belief which is formed or held without those conditions applying, DV will be true of it. Thus,

Doxastic involuntarism (DI) = (EI'') or (UC) when epistemic agents (patients) believe p or form their belief p .

These are the conditions which compromise voluntariness, so they are also the conditions which undermine doxastic voluntariness. A voluntary belief is just one where it is true to say DV of it. An involuntary belief is just one where it is true to say DI of it. We know what the ignorance and compulsion conditions are that voluntariness *simpliciter* needs to avoid. We now need to understand the *doxastic* ignorance and compulsion conditions.

3.2. *The doxastic ignorance condition*

I have formulated the ignorance condition for involuntariness in Part II. I called it EI''. EI'' functions, typically, as a *moral* or *legal* excuse. Speaking in the doxastic domain now, we are searching for a condition which serves as a *doxastic* excuse. This is simply because we are searching for an ignorance condition which excuses epistemic subjects from (doxastic) blame for their *beliefs*.

In classifying this as a doxastic excuse, I'm not asserting that it won't be morally or legally relevant. It is patently false to rule out moral and legal relevance to the formulation of doxastic excuses. Agents generally do things because of their beliefs so assessing whether

⁵⁸ Hereafter, I will omit the 'that' clause when I mention agents' propositional belief that p . 'Believe p ', 'belief p ', or 'believing p ' are all referring to the propositional belief that p (in either verbal or noun form).

someone is doxastically excused is still material to our moral and legal assessment of the matter (even if it is not identical to determining whether EI''' *simpliciter* applies). Further, doxastic excuses can sometimes be *directly* legally relevant. The Sedition Act 1661 prohibited imagining, inventing, and intending (*inter alia*) the death or destruction of the King. If intention is partially constituted by belief-states – which is at least possible – then this law prohibited a set of beliefs.

With this word of caution in mind, the ignorance condition for DI should be a simple matter of transposition. I'll call the doxastic ignorance condition D-EI'''.

D-EI''': if S believes *p* but she lacks true belief about something which, *but for* that lack of true belief, she *wouldn't* believe *p*, and she is not culpable for that lack of true belief, she believes *p* by reason of ignorance which makes her believing *p* *involuntary*. Thus, S is exculpated from blame for believing *p* by way of excuse.

As a condition of DI, D-EI''' looks awry. It appears this way because of two unsettled problems provoked by the formulation. These are:

(a) It is all very well stating that S needs to not be culpable for lacking a true belief which is indispensable for her believing *p*. But wherein lies S's culpability for lacking true beliefs? Moreover, isn't it the case that S can be excused from believing *p* even though she believes *p* in virtue of a *true* proposition?

(b) D-EI''' doesn't exhaust all cases of belief. Not every time S believes some *p* does she believe *p* in virtue of some other belief *q*, false or not. Thus, formulating the condition in terms of non-culpable *lack of true belief* is misguided because some beliefs are held *not* because there is a true belief in the offing which S fails to have for some non-culpable reason, but because there is no belief in the first place which can play an evidential role for *p*. These sorts of belief are *basic* because they don't rely on other propositions to be justifiably held (Kenny, 1992, p.20). In which case, D-EI''' doesn't incorporate the possibility of involuntary basic belief.

In this section, I will address (a-b). I submit, D-EI'' is, despite appearances, largely correct. It, or at least something in the neighbourhood is right. Putting the issues raised by (a-b) to rest will motivate D-EI''s plausibility.

(a): First let us deal with the second question asked in (a): isn't it the case that S can be excused from believing p even though she believes p in virtue of a *true* proposition? D-EI'' stipulates that one is only exculpated from doxastic blame for p if a 'lack of true belief' was instrumental in getting one to believe p . But isn't it the case that people can believe p in virtue of a *true* belief and yet still be exculpated from doxastic blame for p ?

In answering this we shall make a simple but important limit on the type of epistemic agents subject to D-EI''. One ubiquitous way of acquiring beliefs is when the acquired belief is inferred from a prior belief. The latter is evidence for the former. Whether it actually *is* evidence, considered objectively, matters (epistemically) to our inferential capabilities, but it isn't the only thing that matters. Evidence from the *subject's* point of view affects inferential capabilities. The distinction between objective/subjective evidence corresponds to (but isn't the same as) the distinction between objective/subjective rationality. What it is rational for S to do or believe, considered from her perspective, doesn't necessarily mean that it is all-things-considered objectively rational for S to do or believe this-or-that. It is perfectly subjectively rational for S to serve sesame to Shimon *given* S's prior actions and doxastic states; she's served everyone else, and she rationally assumes Shimon can eat sesame.

It can be pressed that what I really mean here is that S can *justify* her belief or action, but that isn't to say that S complies with rationality. This can be shown by making the cases more extreme such that we wouldn't want to label S as being rational in any sense of the word. For example, if S has a final reason to surprise her friend on her birthday, and she believes the best way of doing that would be to chop off her hand give it to her friend as a present,⁵⁹ it seems very strange to say even subjectively, that S was rational. I admit this objection – S isn't rational here. But it remains true that S herself, if pressed for an explanation for her action, can supply reasons she takes to justify her action.

⁵⁹ Kiesewetter and Gertken (2021, p.274).

Similarly, one can have excellent subjective reasons for believing q supports p without it being the case that q is in fact evidence for p .⁶⁰ A detective may reasonably infer that S is the murderer when false evidence has been planted at the crime scene to implicate S. Being presented with the knife with S's fingerprints makes the detective's belief that S did it rational. It is even possible for someone to recognise that they themselves lack objective evidence or reason for p , and yet have some other kind of subjective evidence or reason for p , which makes their believing p rational. This occurs when S acknowledges there is no objective evidence for p , but in virtue of other subjective commitments (prudential or moral, perhaps) S still rationally believes p .⁶¹ Thus, what counts as rational reason for S to believe p doesn't necessarily count as reason for me to believe p .⁶²

Thus, when inferring p from q , two things matter: the subject's subjective epistemic (or belief) situation, and the evidence as it is, objectively. This means that it is possible to rationally infer p from q when neither p nor q are true, when both are true, or when only one of them is true. But what should we say about someone whose inferential capabilities are so haphazard that they take some proposition to support another when the two have *nothing* to do with one another. For example, imagine that S infers p , that Sunderland will win the league, from the proposition q , that Madrid is the capital of Spain. S also has a belief about her inferential methodology: she believes it produces logically valid results, and she won't be persuaded otherwise. Both her beliefs are obviously independently salutary (especially p !), but they patently bear no inferential relation such that one can validly infer p from q . This is obvious to anyone who can see it. Notice the similarity between this case and the one where S chops off her hand. Here too she is hopelessly irrational, and yet if pressed, S can cite reasons explaining her act. The reasons for S are: q , and that she further believes inferring p from q complies with logical norms of inference.

So, what are we to say about S? She has inferred p from a *true* proposition. (a) asks whether S should be excused. Behind this question is the thought that a subject like S expresses such epistemically deviant methods in her reasoning, that it is likely S has cognitive or logical

⁶⁰ Pace Williamson: 'knowledge, and only knowledge, constitutes evidence' (2009, p.185).

⁶¹ But doesn't this face the wrong kinds of reasons problem? Perhaps. I won't discuss this here. If the reader is inclined to think it does face the problem, then the subject isn't rational but like the hand-chopping case, the subject can *explain* her belief. And that explanation can be available to her without being available to you.

⁶² Rational belief (or acceptance) based off this kind of subjective evidence is Kant's '*glaube*'. See Jay (2014, p.214).

impairments that militates against her reasoning rationally. It's plausible enough that someone who genuinely believes to have made a good inference from q to p in this instance, has woeful epistemic apparatus.⁶³ This being S's reasoning procedure, she is in some respect, rationally incapacitated. We need to make the further assumption, however, that S is rationally incapacitated through no fault of her own. I don't know if there is a named condition 'on the books' which refers to *not* merely sloppy reasoning, but a general epistemic deficiency constituted by the belief in the agent, that starkly invalid forms of reasoning are good *and* they reason accordingly; but such a condition isn't theoretically out of bounds. I contend that if this condition describes S, she is a candidate for an epistemic form of the *objective attitude* (Strawson, 2020, p.115). This means that S is *not* excused from her belief p when p is arrived at via a true belief and via her peculiar inferential patterns. Rather, S is outside the scope of doxastic evaluation altogether, so judgments of doxastic in/voluntariness don't apply to her. S is doxastically *exempted* altogether rather than excused. S doesn't meet minimal standards of eligibility for 'involvement or participation' in doxastic appraisal (ibid.). Therefore, it is true that D-EI''' doesn't capture subjects like S who infer p from a true belief q , but that is a virtue.

One should note that D-EI''' *does* include other subjects who reason defectively in a similar fashion. For not every defective reasoner is rationally incapacitated. All of us have epistemic 'blind-spots' where we are liable to be less sedulous and more careless in our inferences. This phenomenon can be motivated by factors like self-deception and gullibility, especially where subjects have something to gain by carelessly inferring. For instance, an otherwise pristine thinker can be tempted by all sorts of biases to support bad inferences if those biases conserve the integrity of some belief or prejudice to which she is unflinchingly committed. This neither excuses nor exempts her from her beliefs arrived at by dodgy inferences. So, she *is* subject to D-EI'''. Her doxastic exculpation depends on whether she is culpable for her liability to reason like so when it comes to propositions (subjectively) germane to her stolid doxastic

⁶³ Granted, different cultural modes of reasoning and inference exist whereby some conclusions reached in one – shall we say – 'logical milieu' is acceptable in one but not another. At one point in human history, to infer that we must sacrifice a child to the gods from the fact that the gods are angry, which is inferred from the fact that there's a drought, would be logically faultless. Now, not so much. (Although, the logical traces are still apparent here in this chain of inference, unlike how it is for S's inference). I am not denying that there is/has been logical relativity amongst cultures. Neither am I claiming (i) that it is obvious we have *warrant* for our rules of logic like *modus ponens* and (ii) this warrant, if we have it, is *epistemic/evidential* in nature. It may very well be the case our reliance on logical rules can only be substantiated by trusting presuppositions for which we cannot give evidence. Yet, that is insufficient to demonstrate we lack warrant for that reliance. Nevertheless, I want to avoid these issues. And I don't think avoiding them prevents anyone from agreeing that if that is how S conducts her logical, inferential operations, then they are no good.

commitment. The person I imagine S to be is someone who *routinely* and *generally* infers ‘apples from oranges.’ The ordinary subject vulnerable to bias doesn’t have her entire reasoning faculties compromised by rational incapacity.

(a) also asked what makes epistemic agents not culpable for their lack of true belief. Smith, in discussing the moral blameworthiness of agents who have acted ignorantly, usefully enumerates three types of *culpable* ignorance (1983, pp.544-48). S is culpably ignorant if one of these disjuncts apply:

- 1) S deficiently investigated.
- 2) S prevented subsequent inquiry.
- 3) S deficiently inferred.

“She should have realised what she was doing” is a slogan encapsulating the moral force behind the idea of culpable ignorance (*ibid.*, p.547). If none of (1-3) applied, the agent would have realised the wrongfulness of her act. (Not that she would have gained new a new *moral* belief, but she wouldn’t have been ignorant of some non-moral fact that makes a moral difference to the case. For instance, medicating patients with medicine M who have underlying condition C, is morally wrong when M and C interact to produce fatal results. The doctor knows this, so she knows the moral fact. Only, she doesn’t know the non-moral fact that her patient, it just so happens, has C. If she didn’t know about C due to one of (1-3), she is culpably ignorant of a non-moral fact that led to moral disaster).

Can we say the same thing in the doxastic domain? Should S have realised what she’s *believing* (in believing *p*)? Realised what, exactly? Well, the standard epistemic correlates of the ethical ‘right/wrong’ couplet, is ‘true/false’. In the moral case “she should have realised what she was doing” means she should have realised, at the time of acting, that what she did was *wrong*. So, in declaring S should have realised what she was believing (in believing *p*), does it make sense to say that S should have realised *p* is false? Yes. This is precisely what doxastic culpable ignorance amounts to. One cannot both realise *p*’s falsity *and* believe *p* at the same time. Thus, that S is culpably ignorant for believing *p* implies that S *doxastically ought not* to believe *p*.

Similar to voluntariness *simpliciter*, doxastic in/voluntarism comprises subjective and evaluative features. So, what is culpable ignorance for S might not be culpable ignorance for someone else. For instance, the criteria for determining doxastic culpable ignorance depends on the agent's role and their relation to the subject matter. I don't carry the same epistemic burdens and obligations when it comes to matters of quantum physics as the professional physicist. What is deficient investigation for her will be no such thing for me. But, towards another subject-matter to which neither of us bare a special relation, the same criteria might apply. Doxastic culpable ignorance is present whenever one, or more, of (1-3) is present. In these instances, S is *not* exculpated from doxastic blame; she believed *p* voluntarily. Yet, determining whether any of (1-3) applies is a task which can only be done on a case-by-case basis by inspecting the believer's role and whether that role has any relevance to the subject-matter to which the belief belongs.

(b): There are many ways we acquire and justify beliefs, and not all of them are by processes of reasoning. This means, that not all beliefs are *evidentially* justified.⁶⁴ Here, I might be using 'evidence' in a somewhat idiosyncratic way, but it is not without precedent. Kenny lists three kinds of basic belief. I shall mention only two because the last one is contentious and involves discussion irrelevant to our purposes. Beliefs are basic iff they are:

- 'Self-evident or fundamental.'
- 'Evident to the senses or to memory.' (Kenny, p.20).

The important difference we are latching onto by introducing basic beliefs is the difference between believing *p* because it follows (or is thought to follow) from some antecedent proposition and believing *p* *immediately* without needing to have explicit reasons for *p* before one believes it. Beliefs from evidence involve *inter alia*, processes of ratiocination, reasoning, pointing to and individuating reasons for *p*. For instance, Mr. X being shown a photograph of Mr. Y 'displaying undue familiarity with Mrs. X' is evidence for Mr. X to infer that Mrs. X is being unfaithful (Grice, 1957, p.382).

Basic beliefs don't depend on identifying reasons for their justification at the moment of acquisition. When *p* is a basic belief, the reason for *p* is *right there* in *p*'s presentation to the

⁶⁴ 'Basic beliefs' is a term often associated with foundationalists about knowledge. But I don't think accepting basic beliefs in this sense need indicate that one is committed to foundationalism.

subject. For instance, the belief that my neighbours are building is formed right away by my seeing them build. No more is required to justify my belief other than my visual sensory input. The justification is right there in my seeing it.

The problem for D-EI'' – as adverted to in (b) – is that it makes no mention of basic non-evidential beliefs. But if basic beliefs include ones which are self-evident, evident to the senses or memory, and perhaps testimony, then D-EI'' doesn't speak to a great deal of our beliefs. Hence, it is hardly of value as a condition of DI.

My first response is to deny that D-EI'' is harmed by the fact that no basic belief can satisfy the condition, because no basic belief *ought* to satisfy the condition. D-EI'' turns on an absence of true belief for which the subject isn't culpable. Its function is to trace faulty doxastic operations which are done involuntarily e.g., S believing something false (*p*) because S lacks some true belief which otherwise would have prevented S believing *p*. This is what it means to believe *p* by reason of ignorance. And 'reasons' imply *evidential* reasons. Yet, since basic beliefs are evidentially unsupported, there is no 'lack of true belief' which explains S forming her basic belief. In other words, with basic beliefs, there is nothing which can stand for S believing *p* by reason of ignorance. Therefore, if S ever has a basic belief, her doxastic voluntariness will not be impaired by ignorance. As far as the test provided by D-EI'' goes, basic beliefs are voluntary.

Yet, this response can be resisted because if we dig deeper, we do find a belief in the vicinity. A belief, which if not possessed *can* stand in for S believing *p* by reason of ignorance. Take testimony, for example. For testimony to yield basic beliefs, agents must have some *background* belief in the doxastic efficaciousness of testimony. Whether this background belief is empirical or *a priori*,⁶⁵ for testimony to elicit beliefs *just* on the speaker's say-so, agents must begin from the belief that testimony is a good, reliable way to attain beliefs. (Or at least the belief that testimony is *not* a bad, unreliable way to attain beliefs). In other words, there must be a default belief shared by all who participate in testimony, which is favourable to the doxastic trustworthiness of testimony. Adler labels this default the '*positive bias*: that one ought simply to accept a speaker's testimony unless one has special reason against doing so' (2006, p.142). One cannot seriously doubt this belief without calling into question the propriety of testimony

⁶⁵ See Moran (2018, pp.39-42) and Adler (2006, pp.144-153).

as an epistemic practice. If S believes contrary to the default position, that testimony is *systematically* unreliable as a method for forming beliefs, then S cannot acquire testimonial basic beliefs. S might yet believe some proposition that was told to her, but she won't be believing it in virtue of its assertoric illocutionary force. Instead, S's belief will be parasitic on evidential reasons *independent* from the testimony. In the same vein, if S seriously doubts the capacity for sensory faculties to deliver reliable information – not merely that S knows she is short-sighted – but that there is something *systematically* epistemically deficient about information derived from the senses, then S's senses cannot provide S with basic beliefs. S must have a background, default belief that *ceteris paribus*, sensory input is sufficiently epistemically reliable to acquire beliefs from its content. Therefore, even basic belief *p* is underpinned by a background default belief. It follows S can believe *p* by reason of ignorance because she can believe *p* whilst lacking the true belief that, say, testimony is generally epistemically reliable.⁶⁶

However, it *doesn't* follow that S can believe *p* by reason of ignorance if she lacks a true background belief. The reason is elucidated in the paragraph above. If S really does lack a true default belief about testimony, she will *not* come to believe testimonial basic beliefs! Recall that the potency of testimony to generate basic beliefs in S depends on S's background, default belief in favour of testimony's epistemic reliability. It won't work as a provider of basic beliefs without it. Ergo, she won't believe (basic belief) *p*. Ergo, D-EI''' doesn't cover cases of basic belief but that is no harm to D-EI''' as a condition of DI.

In summation, D-EI''' serves well as a condition for DI. Furthermore, most of our beliefs are *not* acquired or held by reason of ignorance. Therefore, DV is *pro tanto* true to the extent our beliefs fail to satisfy D-EI'''.

3.3. *The doxastic compulsion condition: On being 'compelled to believe'*

The major thrust for DI springs from the idea that belief is somehow compelled. There are variations of this kind of view. I identified three different types of DI we can find in Aristotle in Part I. They are conceptual DI, psychological DI, and normative DI. Philosophers holding

⁶⁶ I'm assuming plenty in concession to the objector: the truth of these background beliefs, that they *are* beliefs as opposed to dispositions or doxastic practices, and the fact that one cannot believe not-*p* (testimony's unreliability) shows that one must believe *p* (testimony's reliability) when in fact, it shows no such thing.

one of these views might not put it exactly like this, but these three views are united by a commitment to the view that belief is compelled, albeit in different ways and to varying degrees. And thus, DI is true. Since this is the major drive for DI, and the claim that belief is compelled carries significant intuitive appeal, it is crucial I investigate whether belief really is compelled, in the sense germane to involuntarism. In these upcoming sections I will argue that belief is *not* compelled in such a way as to negate voluntariness. Therefore, DI is false and DV is true. To begin, I will quickly dispense with normative DI because this thesis has the weakest ability to claim that belief is compelled. I shall then treat conceptual and psychological DI together, showing that they too fail to establish that belief is compelled in a way that vitiates doxastic voluntariness.

3.3.1. Normative DI

A refresher:

Normative DI: belief is subject to some [unstipulated] standard, and one's freedom to believe is constrained by that standard.

The relevant analogy, as McCready-Flora's interpretation clarifies, is to law (*EE* 1235a2-3). Some things are not-up-to-us because we are subject to normative standards. Law is one kind of normative standard, perhaps a normative standard *par excellence* if only because we can see it being put to practice. If you break the law and your crime is discovered, you are liable to punishment. Therefore, I can't do anything I like. There are legal standards in place preventing such behaviour. In this way, the law acts to constrain my freedom.

Normative DI states something similar about belief. There are doxastic standards and one (normatively) *must* obey those standards. Otherwise, one's belief is bad – it hasn't met the standard it must in order for a belief to be held 'lawfully'. Perhaps the standard is something of which an evidentialist would be proud: one must not believe propositions disproportionately to the evidence for them. 'It is wrong, everywhere, always, and for anyone to believe anything upon insufficient evidence' (Clifford, 1999, p.273). For now, ignore Clifford's ethical flourish.⁶⁷ If this is the normative standard of belief, then we are unfree to

⁶⁷ Or, more deliciously put: ignore Clifford's 'robustious pathos' (James, 1897, p.8).

believe anything which violates it. If we do, we are liable to doxastic repercussions e.g., being subject to reactive attitudes blaming us for our bad beliefs. Therefore, normative DI narrows doxastic freedom.

Although it has been said that law is a ‘system of compulsion on the widest scale’, this cannot substantiate the claim that *norms* (which restrict freedom) themselves compel people to do things involuntarily. Recall that freedom and voluntariness aren’t opposites. Thus, any attempt to analyse compulsion in terms of *unfreedom* won’t alone prove that limitations on freedom imply involuntariness. If that were the case, then every time someone broke the law they would have done so involuntarily. That obviously cannot be true. And neither is it true with beliefs. My doxastic freedom might well be curbed to comply with doxastic norms. But that doesn’t show that my beliefs are involuntary because they are compelled. Therefore, whilst I concur that beliefs are indeed subject to some normative standard, I reject that that shows DI’s truth. This is because being subject to normative standards doesn’t imply involuntariness.

3.3.2. *Conceptual and psychological DI*

A reminder:

Conceptual DI (C-DI): the nature of belief makes doxastic voluntarism a conceptual impossibility.

Psychological DI (P-DI): ordinary human beings (creatures like us) lack the psychological capacities to believe voluntarily.

The norm I stipulated for normative DI mirrors the deontological conception of epistemic justification. The deontological conception maintains that epistemic justification issues prohibitions to epistemic agents. There are principles, rules, or standards which are forbidden to transgress if we are to achieve justified beliefs. Hence, something’s *justification* hinges on it being *permitted* according to its governing norms (Alston, 1988, p.258). Contravening these standards makes us susceptible reproach and blame. Since the archetypical doxastic concerns are to believe truths and not to believe falsities, we should end up with something akin to the evidentialist principle articulated above: beliefs are epistemically justified just in

case they are permitted by the available evidence⁶⁸ because those are the ones sufficiently likely to be true (ibid; Levy, 2007, p.128). Thus far, thus familiar. But where C-DI and P-DI differ from normative DI, is that according to the former two, it is *never* in our power to (i) ensure our beliefs are justified and (ii) to ensure that we avoid unjustified beliefs. Alston predicates the suitability of the deontological conception of epistemic justification on whether we wield voluntary control over beliefs.

Now this conception of epistemic justification is viable only if beliefs are sufficiently under voluntary control to render such concepts as *requirement, permission, obligation, reproach, and blame* applicable to them. By the time-honoured principle that "Ought implies can", one can be obliged to do A only if one has an effective choice as to whether to do A. It is equally obvious that it makes no sense to speak of S's being permitted or forbidden to do A if S lacks an effective choice as to whether to do A. And it seems even more obvious, if possible, that S cannot be rightly blamed for doing (not doing) A if S is incapable of effectively deciding whether or not to do A. Therefore the most fundamental issue raised...is as to whether belief is under voluntary control (ibid., p.259).

And the thought driving most involuntarists is that beliefs *aren't* under voluntary control because beliefs compel. Indeed, we are all familiar with the locution "I'm compelled to believe", when it seems like we have no choice in the matter but to believe *p*. One is often tempted to say this when the evidence points so strongly in favour of *p*; when it is 'barn-door' obvious that *p*. How can one possibly override that irresistible urge to commit doxastically to *p* by choosing to believe not-*p*, or by having the control to refrain from believing *p*, or by believing not-*p* at will? For some philosophers (e.g., Alston), that we don't enjoy such control over our beliefs is reason for repudiating the deontological conception of epistemic justification. In other words, it is a repudiation of the ethics of belief because epistemic reactive attitudes and evaluations make little sense if people aren't in control of their doxastic justification. Not all involuntarists agree with this conclusion, however.⁶⁹ Nevertheless, many involuntarists are involuntarists because of belief's alleged compulsive nature. Here are but a few examples.

⁶⁸ Including the 'non-evidential' reasons to believe present in basic beliefs.

⁶⁹ E.g., Adler (pp.64-7); and Feldman disagrees because he argues doxastic oughts don't imply doxastic cans (2001, p.88). And Adams argues more generally that we are responsible for our states of mind even though they aren't under voluntary control (1985, pp.6-11).

Williams is the most famous proponent of C-DI (1973, p.148). Even if we managed to drastically alter our psychology or transform consciousness so that we could do things at will whilst being completely unaware of what we're doing, something about the concept of belief precludes doing it at will. Belief aims at truth. If we could believe at will, then we could believe anything independent of all truth considerations (Feldman, p.80). But any mental state which accomplishes this would fail to be belief by definition because to *believe* is to do so on the basis of truth considerations.⁷⁰ Being conceptually impossible, one is compelled to believe *p* if one takes the evidence to be for *p*. One is compelled to disbelieve *p* if one takes the evidence to be against *p*.

Owens explicitly denies C-DI – ‘In principle, the will could directly cause anything at all’ – yet he denies we have direct control over beliefs (2000, p.80). Elsewhere, Owens again seems to invoke psychological reasons for denying that we can control beliefs (2002, pp.394f). By some measure then, Owens subscribes to P-DI along with Alston who ‘cannot see any sufficient reasons for the stronger claim’, C-DI (ibid., p.263). Smith also holds that our inability to exercise control over beliefs and other nondoxastic attitudes is because of our psychological constitution: we just don’t have the right psychology to choose or control these attitudes (2005, p.265).

On one register, P-DI seems to be a thesis about intellectual *skill*. Controlling one’s beliefs is a very difficult skill involving an excellence so great, no one can master it. It is a bit like the skill required to be able to recite all of *War and Peace* by heart after only one read. This is certainly no conceptual impossibility. But creatures like us are unfortunate enough to never be able to obtain that skill. In another register, P-DI states something slightly stronger. It is part of our psychological *constitution* that we cannot directly control beliefs. There is no skill – for us – which could ever correspond to directly controlling beliefs. Perhaps a creature like God could do it. But for us, with our psychology, it is impossible. But this psychological impossibility is still a *contingent* impossibility. There is nothing metaphysically necessary in having the psychology we do. I think that most advocates of P-DI have the latter iteration in mind. Nevertheless, whichever way we lack the psychological ability to control our beliefs, we *cannot*

⁷⁰ This is a little rushed, but it conveys the core of the argument well enough. For criticism of Williams’ argument, see Winters (1979). She argues that the most Williams shows is that *sustaining* beliefs cannot be done at will but *acquiring* beliefs might yet be basic actions done at will (pp.253-6).

directly control them, nor can we will them. Thus, *ex hypothesi*, beliefs compel. We *must* believe *p* if we think the evidence is for *p*, and we cannot exercise control or our will over them.

3.3.3. *Voluntarism without the will: Un-willing beliefs*

Some motivations for DI stems from belief's putative compelling nature, and this is *frequently* couched in terms of *voluntary control* and believing *at will*. Given our findings in Part II, we ought to be confident now that this is a mistake. It is evident that the doxastic involuntarist/voluntarist debate has also been contaminated by conflating activity, acts of will, or volitions with *doxastic* activity. So it too has been infected by the 'modern theory of the will'. If we want to understand *voluntary* belief, we need to look away from the agential dimension of human conduct and consider the evaluative realm. This means that we need to prevent conationist notions like 'believing *p* at will' from contaminating the discourse about doxastic voluntarism. In other words, we need to *un-will the will* from our understanding doxastic voluntarism and involuntarism.

Additionally,⁷¹ we ought to eliminate talk of will from DI/DV if 'at will' means doing something directly, whenever one wants, or easily. Like when I can raise my right arm whenever I want, with no difficulty whatsoever. But suppose someone has paralysis in their right arm and so whenever they want to raise their right arm, they must use their left arm to indirectly bring it about that they raise their right arm. That is, actions the like of which, we must *make ourselves to bring about indirectly*. The same temptation to lead someone to identify doing 'x at will' with the former kind of action, is liable to lead someone to think that the latter types of action are *not* done at will. Precisely because the possibilities for this person to raise her right arm are indirect, not always available, and not as easy as it is for me.

This is the notion of doing 'x at will' Bernard Williams has in mind when he discusses the differences and similarities between blushing and believing (*ibid.*). Williams refuses to simply say 'you can blush at will' by making yourself bring about your own blushing because that involves blushing 'by a route'. In other words, by indirect, difficult means. But this notion is wrong. Think back again to the person with a paralysed right arm. Just because she has to mobilise her left arm in order to raise her right arm doesn't diminish the 'at will-ness' of her

⁷¹ The next two paragraphs are shaped by Hieronymi (2009), including the soup example (p.159).

raising her right arm. It just makes it a little bit less convenient. Similarly, I can't make soup as easily or as directly as I can raise my right arm. But from that fact we don't thereby derive that I *didn't* make soup at will. Contra Williams, doing x and making yourself via some protracted means bring about x , is not the difference between doing something at will and not doing something at will. This is important because it evinces that *none* of these distinctions settle the question of DI/DV. If x can be done at will and directly, that isn't to say x is done either voluntarily or involuntarily. If x can be done at will and *indirectly*, that too isn't to say x is done voluntarily or not. If x can be done *not* at will and directly, ditto. Finally, if x can be done *not* at will and *indirectly*, ditto. This sense of 'at will' which we just distinguished from acting directly, is kindred to the willingness with which we ϕ when we *don't* ϕ reluctantly (see 2.4.4.). It tells us nothing about whether ϕ was compelled and therefore, it tells us nothing about whether ϕ was voluntary. Hence, we ought to un-will notions of will and 'at will' from the question about voluntary belief.

3.3.4. *Evidentialism and compulsion again*

In Part 2.3., I introduced a distinction between something's being compelled and something being *under* compulsion. I argued that only the latter are compelling factors which diminish voluntariness. The sort of things which are compelled is the beating of my heart, the falling of my pen when I drop it, and the motion of billiard balls when struck with enough force. The sort of things which are UC are human agents who are yoked with an unwelcome force that makes them do something contrary to their desires and preferences. There is a gap between what the compelling force makes hypothetically necessary, and what the subject UC wants.

In agreement with the involuntarists, beliefs are usually things which have a feel of force to them. To simplify our lives, let's focus only on beliefs for which we have overwhelming evidence or basic beliefs which are self-evident. I believe that I am sat in my bedroom writing on my laptop. I believe two and two is four. I believe that Australia exists.⁷² Indeed, I *know* all these things. And it really does seem impossible for me to believe otherwise on my own accord. There is no voluntarist who could plausibly deny this. The case can be adjusted to make a denial slightly more plausible e.g., I take a magic pill which instantly makes me believe the negation of these propositions. But even then, it wouldn't be *me*, on my *own accord*

⁷² Kenny's example (ibid., pp.14ff).

achieving those beliefs and then those beliefs would feel forced. The reason is because of belief's orientation to truth. Beliefs have representational content because they aim to represent reality. This means that for any one of any subjects' belief, that belief is formed or held in virtue of truth-considerations.⁷³ Once we see that there is sufficient reason for believing p , we believe p . The truth-considerations cannot be substituted for any other kind of reason. Desire, for example, is the wrong kind of reason for belief. You can *incentivise* me to believe something I believe is false. You might try to persuade me that it is morally better to believe p than to not, you might point out all the practical benefits of believing p , or you might even offer me £10,000 to believe p . No matter the intensity of my desire, I cannot believe something I take to be false, just like that.

However, it is possible to alter my beliefs *indirectly* by interfering with my noetic structure. But this takes a protracted epistemic process.⁷⁴ I'd have to be selective with the evidence I allow to present itself forth to me, surround myself with believers, commit acts of self-deception or at least self-suggestion, and so on. By these means, I *might* be able to come to believe Australia doesn't exist or that the Earth is flat. Nevertheless, this is only accomplished – if it is accomplishable at all – by manipulating my truth-considerations for p . If I have managed to overturn my belief p in favour of $\neg(p)$, it is only because I now take reasons to believe $\neg(p)$ to be sufficient. My desire to believe $\neg(p)$ is never the effective cause of my belief $\neg(p)$. Truth-considerations are the only right doxastic reasons; never desire. Thus, it seems that beliefs are non-optional for subjects.

All this is redolent of *evidentialism*. The fundamental principle, alluded to earlier, is that one *ought* only to believe p based on evidence (or reasons) germane to the truth of p . So, it is a normative and not a descriptive thesis. However, motivations for this norm can be grounded in something *descriptive*. This can be done in a way either reflective of C-DI or P-DI. Adler's appellation for the claim outlined above – that one's beliefs are believed on the basis of the

⁷³ This assumption only becomes controvertible in cases more complex than self-evident beliefs or ones for which we have overwhelming evidence. I might believe p not for truth-considerations, but because I want to investigate whether p is true (or perhaps I want to *disprove* p). I do this when I argue *reductio ad absurdum*: to see if a contradiction is derivable from p , I need to use p as a premise in my argument. Although, this is probably more aptly characterised as an instance of *nondoxastic acceptance*. See Cohen (1992, p.13). Moral fictionalists might believe moral propositions not for truth-considerations, but for prudential reasons. Then again, this might too be best characterised as nondoxastic acceptance.

⁷⁴ Processes of exercising 'long range control' over beliefs (Alston, p.275).

subject taking herself to have sufficient reason for beliefs – is the ‘*subjective principle of sufficient reason*’ (SPSR) (p.26). How does Adler establish this principle as fact?

The main reason to believe that the SPSR is a fact is that we find ourselves *compelled*⁷⁵ to follow it... That we do follow it is then a reflection of our grasping the demands of belief, not merely a curious psychological truth about us. (Adler, p.27)

For Adler, evidentialism is a norm grounded by a descriptive fact: by a fact about the conceptual nature of belief. Alternatively, evidentialism qua norm can be motivated by our psychological constitution: it just so happens that our ‘belief-forming faculties are... set up to be sensitive to evidence’ (Chignell, 2018, §5.3). Therefore, normative evidentialism reflects descriptive facts, either conceptual or psychological, and these facts display belief’s non-optionality.

However, I can grant the descriptive truth of the conceptual and/or psychological motivations for normative evidentialism whilst resisting the claim that beliefs are *involuntary*. What do beliefs (forming them or having them) resemble more? Are they, in certain respects, like my beating heart, the motion of a struck billiard ball, and a vending machine expelling its goods? Or are they, in certain respects, like being subject to a coercive threat? I answer that beliefs resemble the first group of examples, and they don’t resemble being subject to coercive threats.

Granted, when I believe *p* it is non-optional. But it is only non-optional in the same sense it is non-optional for my heart to beat and a vending machine to spew out KitKats when the right code is punched in. Now, this doesn’t mean that vending machines or beating hearts perform their operations voluntarily. They don’t because they’re neither doxastic nor evaluative creatures. Thus, the conditions of EI”” and UC aren’t *ab initio* applicable to them.

We, however, are legitimate candidates for voluntariness and involuntariness. *We’re* not like vending machines, but beliefs and/or the epistemic processes which we own *are* like vending machines. They are akin only because we are we are *compelled* to believe by evidence and truth-considerations when we deem them sufficient, and the machine is compelled to do

⁷⁵ My emphasis.

something when there is sufficient 'reason' i.e., I enter-in a code. That is the crucial parallel to acknowledge. Either the conceptual nature of belief makes it so, or the way our epistemic apparatus happens to be makes it so. But as argued in 2.3., that doesn't disclose the fact that we are *under compulsion* to believe.

The vending machine analogy is useful. It won't eject any item at any time. The right buttons need to be pressed in the right order to get something. As soon as a right combination in the right order is entered, and the money has been paid, then the vending machine has adequate or sufficient 'reason' to provide its goods. The vending machine doesn't have an option. It must execute the right operations to give me my KitKat. But it isn't UC to do so. The vending machine might malfunction and fail to give me my KitKat. This isn't proof that the KitKat has options it prefers, but that faulty mechanisms can generate substandard or unexpected results. The combination of buttons and money paid stands for doxastic truth-considerations and evidence. Once we deem our reasons for p to be adequate, we *must* believe p . But we're not UC to do so. It's just how our 'epistemic machinery' works. Likewise, if it is possible for someone to *not* believe p when they have sufficient subjective reason to believe p this doesn't prove that the subject has doxastic options it would prefer. It merely exhibits a malfunction within the subject's epistemic machinery. (However, that this could ever happen is extremely doubtful. For S to say she has sufficient evidence for p and yet she doesn't believe p is for S to assert a paradoxical Moore-statement: " p , but I don't believe p ". This is because of the *transparency* of belief, which is a condition about *how* to answer two distinct questions (Moran, 2001, p.62). S determines the question of whether she believes p , by determining the world-directed question of 'whether p ?' i.e., whether S takes herself to possess sufficient reason for p . Whether I believe there will be a third world war is answered by my thinking about the evidence for the proposition that there will be a third world war.⁷⁶ Hence, two distinct questions are treated equivalently. They are transparent. Consequently, if S fails to believe p despite claiming sufficient reason for p , she violates this transparency condition which seems impossible). Therefore, beliefs compel but they do not place subjects UC.

Crucially, this is true about how we *routinely* acquire beliefs and believe. If everything runs smoothly, then we *always* believe what we must. And what we must believe is governed by conceptual and/or psychological facts about belief. That subjects non-optionally believe p so

⁷⁶ Evans' example, quoted in (Soteriou, 2005, p.102).

long as they have sufficient subjective reason for p . Hence, the vast majority of our beliefs are compelled in this way. It follows that the vast majority of our beliefs are acquired and held *without* the compulsion condition for involuntariness applying. Therefore, not only is DV theoretically possible, not only is it practically possible, but DV is true *most* of the time of *most* of our beliefs.

3.4. *Acts and states – agency, patiency*

I've argued that DV is generally true. Some philosophers who think about doxastic in/voluntarism and doxastic responsibility are really worried about salvaging an *agency* that we exercise over our beliefs. Beliefs appear to be passive states rather than mental acts. Moreover, they seem to be states which happen *to* us rather than us bringing them about. If what I have argued up till now is roughly correct, then the claim that we are *patients* to our beliefs is plausible. All subjects do is examine the truth-considerations and evidence, and if it is enough for belief, a belief is formed. Boyle's (2011) argument is admirably creative. Boyle attempts to maintain belief's state-like appearance whilst accounting for doxastic agency as the activity of *being* in a certain state.

Ultimately, I think Boyle is wrong, but I won't contend that here. My purpose in this section is to argue that according to my view of DI, the problems of acts/states and agency/being patients isn't a quagmire we need enter. Voluntariness is a quality of *both* acts and states, and it is attributable to *both* agency and patiency so long as the object of our attribution meets the minimal condition for voluntariness (2.8). On my account, DV doesn't imply anything about the active or agential status of belief. Whatever beliefs are, and whatever our agential relation to them (or none), DV is still true. We might passively undergo beliefs as patients, or actively believe as agents. It's a virtue of my theory that it can accommodate advocates of either disjunct. DV, properly understood, is compatible with both doxastic agency and doxastic patiency.

3.5. *An objection – wanting, wishing, and 'outward criteria'*

I've conceded that beliefs are compelled but denied that we are doxastically UC. UC is defined as being in a condition such that there is tension between what you want and what must be done. But it is a familiar phenomenon for subjects to believe p when they don't want

to! Mothers might not want to believe their sons are criminals, but they nevertheless believe it when they have sufficient subjective reason for the proposition. Lovers don't want to believe their partners are unfaithful, but they will if they have sufficient subjective reason for the proposition. Therefore, in these cases, subjects are doxastically UC. They have a preference to not believe p . But their subjective reasons for p acts as a compelling force to make them believe p . So, there is a tension between what subjects want and what must be done.

Is it true that these subjects want to not believe p ? What is the nature of their desire?⁷⁷ Let us attend to, briefly, the nature of desire in general. Desires are *dispositions*. Like all dispositions, desires can be *manifested*. Desires are manifested in two ways (Hyman, pp.107ff).

- 1) Goal-directed behaviour. Behaviour aimed at obtaining the object of desire or doing what it is a desire to do.
- 2) Affective reactions such as gladness and relief (or disappointment) if the desire is satisfied (or frustrated).

There are other ways desires express themselves, but only their manifestations are constitutive of their identity (ibid.). This is the difference between feelings' *signs* and *symptoms* and their *criterion*. I know that someone has a toothache when she holds her cheek because I do the same when I have toothaches (Wittgenstein, 1958, p.24). This is merely a sign, not a constitutive definition of toothaches. Its criterion, I suppose, is the aching pain itself. My desire for X can be expressed via signs and symptoms e.g., my salivating is a symptom of my desire to eat steak. Only eating steak, or actions taken as means to eating steak are manifestations/criteria of the desire.⁷⁸ It is what *identifies* the desire. Though I indicate my desire by salivating when thinking of steak, my desire is no desire to salivate; it is a desire to eat steak. This doesn't mean that to have a desire one must be able to express the desire through action. However, it *does* mean for any desire D, there must be a way to *manifest* D. If some feeling or longing has no means of ever being manifested, it is no desire. 'An 'inner process' stands in need of an outward criteria' (Wittgenstein, 2009, §580). This is what separates mental processes like desire, from indolent feelings or thoughts like wishes (Hyman, ibid.). I can desire things only if the disposition can be manifested, only if it has the possibility

⁷⁷ 'Desire' used interchangeably with 'want'.

⁷⁸ The example is Hyman's (p.108).

of a criterion being outwardly expressed. Wishing, however, isn't so constrained. I can *wish* to fly, but I cannot desire it. I can wish the impossible, but only want the possible.

My objector claims it is well-known that people believe propositions when they don't want to. Do they *want* to not believe it, or do they *wish* it? Subjects can want to not have their evidential perspective on p be sufficient for believing p . This is a dispositional desire because there is a way to manifest it. Its outward criterion is exhibited by S simply judging the evidence to be insufficient to warrant believing p . But this is not the case at hand. In wanting to believe her lover is faithful despite her belief that he is unfaithful, we are not presuming that S radically transforms her evidential perspective regarding p . Because if that were the case, S would simply *not* believe p . The problem that the objection attempts to highlight is that S *retains* her evidential perspective on p – so that she still thinks there is sufficient reason to believe p – and yet doesn't want to believe p . But *this* is no want. Why? Because there is nothing that could possibly manifest a disposition to both see the facts as corroborating p and yet not believe p . That would infringe either a conceptual impossibility or a psychological impossibility. It is a yearning for S to not have her doxastic mechanisms sensitive to truth-considerations and evidence. This means that the object of the feeling has no manifestable outward criterion. It can't be done, so it cannot be a desire or want to not believe p against S's own take on the evidence. S can wish for it, however.

Thus, the 'familiar phenomenon' the objection relies on ought to be modified. People don't want to not believe in ways incommensurate with their subjective reasons for believing p , though they might wish for it. But wishes which are in tension with what must be done isn't enough for UC to apply. Voluntariness is incompatible with forces which make S do things she doesn't *want* to do. Voluntariness *is* compatible with forces which make S do things she doesn't *wish* to do, because subjects having wishes inconsistent with what must be done is insufficient for UC. Therefore, the objection doesn't successfully militate against my thesis that DV is routinely true.

3.6. Conclusion

In Part 2 we learnt that voluntariness is a conceptual tool for legal and ethical evaluation. Its dominion primarily stretches over the ethical domain of human conduct. Our beliefs are also subject to evaluative practices, so oughtn't we be able to determine when they are had

voluntarily or involuntarily? If doxastic in/voluntarism should say anything about *voluntariness*, we need to un-will notions of will from our understanding of the theses. I have accomplished this in this Part 3.

DV ought to be understood as a thesis about the absence of doxastic exculpatory conditions when acquiring or holding beliefs. I argued contra epistemological orthodoxy, DV is true, and routinely true. Most of the time, we form and hold beliefs without being doxastically EI'' and without being under compulsion. This is true even though beliefs compel. They compel by dint of conceptual or psychological impossibilities. Lastly, I dealt with an important objection concerning the feeling of not wanting to believe something one believes. I argued that such a case is properly characterised as a case of wishing for something, not wanting. And only wanting what cannot be done, not wishing what cannot be done, is inconsistent with DV.

Conclusion

Compared to its opponent, doxastic involuntarism seems a natural and sensible view. Indeed, subscribing to voluntarism evinces an unenviable lack of philosophical caution. But if my arguments are sound, I have shown that there is good reason to turn this presumption on its head. Involuntarism *seems* the way to go, but only because epistemology has been infected by an error whose home is in the philosophy of mind and action. If voluntary acts are thought to be the movements of an agent in virtue of some precipitating act of will, volition, or conation, then voluntary belief easily becomes associated with an ability to believe at will. *Un-willing* beliefs, however, has the salutatory effect of illuminating the relation between beliefs and voluntariness, properly understood. I argued it is properly understood as an evaluative concept designed to identify when doxastic and evaluative creatures are *prima facie*, suitably susceptible to ethical and legal assessment. Voluntariness is destroyed by the presence of either the ignorance or the compulsion condition. This typically provides grounds for *exculpation*. When agents or patients are *excused* from guilt or full blame.

Part 3 articulated the corresponding conditions for doxastic involuntariness: the doxastic ignorance condition and the doxastic compulsion condition. Doxastic voluntarism, therefore, is true just in case those two conditions are absent in cases of forming and holding beliefs. This is true regardless of whether we our doxastic *agents* or *patients*. I argued that the way we acquire beliefs and hold them don't satisfy the conditions for doxastic involuntarism. The fact that beliefs *compel* is insufficient to substantiate that we are *under* their compulsion. Therefore, I conclude that doxastic voluntarism is routinely true.

Bibliography

- Adams, R. M. (1985). Involuntary sins. *The Philosophical Review*, 94(1), 3.
- Adler, J. E. (2006). *Belief's own ethics* (1. paperback ed). MIT Press.
- Alston, W. P. (1988). The deontological conception of epistemic justification. *Philosophical Perspectives*, 2, 257.
- Alvarez, M. (2010). Thomas Reid. In T. O'Connor & C. Sandis (Eds.), *A Companion to the Philosophy of Action* (pp. 505–512). Wiley-Blackwell.
- Alvarez, M. (2016). The concept of voluntariness. *Jurisprudence*, 7(3), 665–671.
- Anscombe, G. E. M. (2000). *Intention* (2nd ed). Harvard University Press.
- Aristotle. (2009a). The complete works of Aristotle. 2. In J. Barnes (Ed.), *Metaphysics* (6. print., with corr., [Nachdr.]). Princeton Univ. Pr.
- Aristotle. (2009b). *The Nicomachean Ethics* (W. D. Ross, Trans.). Oxford University Press.
- Aristotle. (2011). *The Eudemian Ethics* (A. Kenny, Ed. & Trans.). Oxford University Press.
- Aristotle. (2016). *De Anima* (C. Shields, Ed. & Trans.). Oxford University Press.
- Austin, J. L. (1962). *Sense and sensibilia* (G. J. Warnock, Ed.). Oxford Univ. Press.
- Austin, J. L. (1979). A plea for excuses. In J. O. Urmson & G. J. Warnock (Eds.), *Philosophical Papers* (3rd ed., pp. 175–203). Oxford University Press.
- Baltzly, D. (1997). Knowledge and belief in Republic V. *Archiv Für Geschichte Der Philosophie*, 79(3).
- Baron, M. (2017). Justification, excuse, and the exculpatory power of ignorance. In R. Peels (Ed.), *Perspectives on ignorance from moral and social philosophy* (pp. 53–76). Routledge.
- Bentham, J. (1789). *An introduction to the principles of morals and legislation*. T. Payne and Son.
- Black, H. C. (1999). *Black's law dictionary* (B. A. Garner, Ed.; 7. ed). West Group.
- Boyle, M. (2011). 'Making up your mind' and the activity of reason. *Philosopher's Imprint*, 11(17).
- Brown, L. (2009). Explanatory notes. In *The Nicomachean Ethics*. Oxford University Press.
- Chignell, A. (2018). The ethics of belief. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.
<https://plato.stanford.edu/archives/spr2018/entries/ethics-belief/>.
- Clifford, W. (1999). The ethics of belief. In E. Stump & M. J. Murray (Eds.), *Philosophy of religion: The big questions* (pp. 269–273). Blackwell Publishers.
- Cohen, L. J. (1992). *An essay on belief and acceptance*. Clarendon Press.
- Daube, D. (2008). *The deed and the doer in the Bible* (C. M. Carmichael, Ed.). Templeton Foundation Press.

- Davidson, D. (2001). Actions, reasons, and causes. In *Essays on actions and events* (2nd ed, pp. 3–19). Clarendon Press ; Oxford University Press.
- Descartes, R. (1988). *Descartes: Selected philosophical writings* (J. Cottingham, R. Stoothoff, D. Murdoch, & A. Kenny, Eds.; 1st ed.). Cambridge University Press.
- Descartes, R. (1998). *Meditations and other metaphysical writings* (D. M. Clarke, Trans.). Penguin.
- Descartes, R. (2017). *Descartes: Meditations on first philosophy: with selections from the objections and replies* (J. Cottingham, Ed. & Trans.; 2nd ed.). Cambridge University Press.
- Desmedt et al, O. (2022). What do measures of self-report interoception measure? Insights from a systematic review, latent factor analysis, and network approach. *Biological Psychology*, 169, 108289. <https://doi.org/10.1016/j.biopsycho.2022.108289>.
- Duff, R. A. (1981). Recklessness and rape. *The Liverpool Law Review*, 3(2), 49–64.
- Feldman, R. (2001). Voluntary belief and epistemic evaluation. In M. Steup (Ed.), *Knowledge, truth, and duty: Essays on epistemic justification, responsibility, and virtue* (pp. 77–92). Oxford University Press.
- Fine, G. (1978). Knowledge and belief in Republic V. *Archiv Für Geschichte Der Philosophie*, 60(2).
- Frankfurt, H. G. (1998a). Coercion and moral responsibility. In *The importance of what we care about: Philosophical essays* (pp. 26–46). Cambridge University Press.
- Frankfurt, H. G. (1998b). Freedom of the will and the concept of a person. In *The importance of what we care about: Philosophical essays* (pp. 11–25). Cambridge University Press.
- Frankfurt, H. G. (1998c). The importance of what we care about. In *The importance of what we care about: Philosophical essays* (pp. 80–94). Cambridge University Press.
- Frankfurt, H. G. (1998d). The problem of action. In *The importance of what we care about: Philosophical essays* (pp. 69–79). Cambridge University Press.
- Grice, H. P. (1957). Meaning. *The Philosophical Review*, 66(3), 377.
- Guthrie, W. K. C. (1956). Note on the experiment with the slave. In *Plato's Protagoras and Meno* (pp. 107–114). Penguin.
- Hart, H. L. A. (2008a). Postscript: Responsibility and retribution. In *Punishment and responsibility: Essays in the philosophy of law* (2nd ed, pp. 210–237). Oxford university press.
- Hart, H. L. A. (2008b). Prolegomenon to the principles of punishment. In *Punishment and responsibility: Essays in the philosophy of law* (2nd ed, pp. 1–27). Oxford university press.
- Hart, H. L. A., & Honoré, T. (1985). *Causation in the law* (2nd ed). Clarendon Press; ;Oxford University Press.
- Hieronimi, P. (2009). Believing at will. *Canadian Journal of Philosophy Supplementary Volume*, 35, 149–187.

- Hieronymi, P. (2020). *Freedom, resentment, and the metaphysics of morals*. Princeton University Press.
- Hieronymi, P. (2022). What is a will? In U. Maoz & W. Sinnott-Armstrong (Eds.), *Free will: Philosophers and neuroscientists in conversation* (pp. 13–20). Oxford University Press.
- Hobbes, T. (1998). *Leviathan* (J. C. A. Gaskin, Ed.). Oxford University Press.
- Hobbes, T., & Bramhall, J. (1999). *Hobbes and Bramhall on liberty and necessity* (V. Chappell, Ed.). Cambridge University Press.
- Hornsby, J. (1980). *Actions*. Routledge and Kegan Paul.
- Hyman, J. (2015). *Action, knowledge, and will*. Oxford university press.
- Hyman, J. (2016). Voluntariness and intention. *Jurisprudence*, 7(3), 692–709.
- James, W. (1897). The will to believe. In *The will to believe and other essays in popular philosophy* (pp. 1–31). Longmans, Green and Co.
- Jay, C. (2014). The Kantian moral hazard argument for religious fictionalism. *International Journal for Philosophy of Religion*, 75(3), 207–232.
- Kenny, A. (1982). Duress per minas as a defence to crime II. *Law and Philosophy*, 1(2), 197–205.
- Kenny, A. (1992). *The metaphysics of mind*. Oxford university press. (Original work published 1989)
- Kenny, A. (1992). *What is faith? Essays in the philosophy of religion*. Oxford University Press.
- Kiesewetter, B., & Gertken, J. (2021). How do reasons transmit to non-necessary means? *Australasian Journal of Philosophy*, 99(2), 271–285.
- Le Morvan, P., & Peels, R. (2016). The nature of ignorance: Two views. In M. Blaauw & R. Peels (Eds.), *The Epistemic Dimensions of Ignorance* (pp. 12–32). Cambridge University Press.
- Levy, N. (2007). Doxastic responsibility. *Synthese*, 155(1), 127–155.
- Lewinsohn, J. (2006). Philosophy of halakha: The case of intentional action. *Torah U-Madda Journal*, 14, 97–136.
- Locke, J. (1990). *An essay concerning human understanding* (J. W. Yolton, Ed.). Dent & Sons Ltd.
- McCready-Flora, I. C. (2013). Aristotle and the normativity of belief. *Oxford Studies in Ancient Philosophy*, 44, 67–96.
- Melden, A. I. (1956). Action. *The Philosophical Review*, 65(4), 523–541.
- Moran, R. (2001). *Authority and estrangement: An essay on self-knowledge*. Princeton University Press.
- Moran, R. (2018). *The exchange of words: Speech, testimony, and intersubjectivity*. Oxford University Press.
- Moss, J. D. (2021). *Plato's epistemology: Being and seeming* (First edition). Oxford University Press.
- Moss, J., & Schwab, W. (2019). The birth of belief. *Journal of the History of Philosophy*, 57(1), 1–32.
- Nahmanides. (1990). Ramban al ha-Torah. In *Torat Chayim—Sefer Vayikra (Leviticus)*. Mosad ha-Rav Kook.
- Nozick, R. (1997). Coercion. In *Socratic puzzles* (pp. 15–44). Harvard University Press.

- O'Brien, L. (2007). *Self-knowing agents*. Oxford University Press.
- Owens, D. (2000). *Reason without freedom: The problem of epistemic normativity*. Routledge.
- Owens, D. (2002). Epistemic akrasia. *The Monist*, 85(3), 381–397.
- Peels, R. (2010). What is ignorance? *Philosophia*, 38(1), 57–67.
- Pink, T. (2009). Reason, voluntariness, and moral responsibility. In L. O'Brien & M. Soteriou (Eds.), *Mental actions* (pp. 95–119). Oxford University Press.
- Plato. (1956). *Protagoras and Meno* (W. K. C. Guthrie, Trans.). Penguin.
- Plato. (2007). *The Republic* (D. Lee, Trans.; 2nd ed.). Penguin Books Ltd.
- Reid, T. (2010). *Essays on the active powers of man* (K. Haakonssen & J. A. Harris, Eds.). Edinburgh University Press.
- Rosen, G. (2008). Kleinbart the oblivious and other tales of ignorance and responsibility. *The Journal of Philosophy*, 105(10), 591–610.
- Ryle, G. (2000). *The concept of mind* (Reprinted with a new introduction). Penguin Books. (Original work published 1949)
- Shakespeare, W. (2007). The tragedy of Macbeth. In J. Bate & E. Rasmussen (Eds.), *Complete works: The RSC Shakespeare* (pp. 1859–1917). Macmillan.
- Sharon, A. (2016). Action, voluntariness and consent: On John Hyman's *Action, knowledge, and will*. *Jurisprudence*, 7(3), 678–684.
- Smith, A. M. (2005). Responsibility for attitudes: Activity and passivity in mental life. *Ethics*, 115(2), 236–271.
- Smith, A. M. (2015). Responsibility as answerability. *Inquiry*, 58(2), 99–126.
- Smith, H. (1983). Culpable ignorance. *The Philosophical Review*, 92(4), 543.
- Soteriou, M. (2005). Mental action and the epistemology of mind. *Nous*, 39(1), 83–105.
- Steward, H. (2012). *A metaphysics for freedom*. Oxford University Press.
- Strawson, P. F. (2020). Freedom and resentment. In *Freedom, resentment, and the metaphysics of morals*. (pp. 107–133). Princeton University Press.
- Stump, E. (1988). Sanctification, hardening of the heart, and Frankfurt's concept of free will. *The Journal of Philosophy*, 85(8), 395–420.
- Torah Neviim Ketuvim (Torah, Prophets, Writings)*. (2009). Koren.
- Wedin, M. V. (1988). *Mind and imagination in Aristotle*. Yale University Press.
- White, A. R. (1968). Introduction. In A. R. White (Ed.), *The philosophy of action* (pp. 1–18). Oxford University Press.
- Williams, B. (1973). Deciding to believe. In *Problems of the self: Philosophical papers 1956–1972* (1st ed., pp. 136–151). Cambridge University Press.

- Williams, B. (1993). *Shame and necessity*. University of California Press.
- Williams, B. (1995). Voluntary acts and responsible agents. In *Making sense of humanity and other philosophical papers, 1982-1993* (pp. 22–34). Cambridge University Press.
- Williamson, T. (2009). *Knowledge and its limits* (Reprint). Oxford University Press.
- Winters, B. (1979). Believing at will. *The Journal of Philosophy*, 76(5), 243.
- Wittgenstein, L. (1958). *The Blue and Brown Books* (Nachdr.). Blackwell.
- Wittgenstein, L. (1967). *Zettel* (G. E. M. Anscombe & G. H. von Wright, Eds.; G. E. M. Anscombe, Trans.). Basil Blackwell.
- Wittgenstein, L. (2001). *Tractatus logico-philosophicus* (D. Pears & B. McGuinness, Trans.). Routledge.
- Wittgenstein, L. (2009). *Philosophische untersuchungen =: Philosophical investigations* (P. M. S. Hacker & J. Schulte, Eds.; G. E. M. Anscombe, Trans.; Rev. 4th ed). Wiley-Blackwell. (Original work published 1953)
- Wolf, S. (1987). Sanity and the metaphysics of moral responsibility. In F. D. Schoeman (Ed.), *Responsibility, character, and the emotions: New essays in moral psychology* (pp. 46–62). Cambridge University Press.
- Wolf, S. (1990). *Freedom within reason*. Oxford University Press.
- Zimmerman, M. J. (2002). Controlling ignorance: A bitter truth. *Journal of Social Philosophy*, 33(3), 483–490.