Mindlessness

Thinking is overrated: golfers perform best when distracted and under pressure; firefighters make the right calls without a clue as to why; and you are yourself ill advised to look at your steps as you go down the stairs, or to try and remember your pin number before typing it in. Just do it, mindlessly. Both empirical psychologists and the common man have long worked out that thinking is often a bad idea, but philosophers still hang on to an intellectualist picture of human action. This book challenges that picture and calls on philosophers to wake up to the power of mindlessness: it is our habits, skills and conventions that help us cope with a world way too diverse for us to hope to always reinterpret it. The book presents the empirical evidence that has been accumulating over the last few decades and offers a philosophical analysis of mindless phenomena such as habits, skilled activity, automatic actions, emotional and spontaneous reactions and social conventions, arguing that traditional philosophical theories of action should be revised to do justice to this forgotten but important part of our lives: when we act mindlessly, we are free and fully rational even though we neither deliberate nor are aware of what we are doing.

Ezio Di Nucci is Assistant Professor of Philosophy at Universität Duisburg-Essen, having previously taught at the University of Edinburgh (where he received his PhD in 2008), the University of Stirling and University College Dublin. On top of many journal articles, he has published two other books: *Content, Consciousness, and Perception* (2006, with Conor McHugh) and *Mind Out of Action* (2008).

Mindlessness

By

Ezio Di Nucci

CAMBRIDGE SCHOLARS

PUBLISHING

Mindlessness, by Ezio Di Nucci

This book first published 2013

Cambridge Scholars Publishing

12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-4438-4658-9, ISBN (13): 978-1-4438-4658-5

a Usa,
che questo libro l'ha già fatta ridere

… und schreib getrost
“Im Anfang war die Tat”

(Wittgenstein, *On Certainty*, §402,
originally in Goethe’s *Faust*)

TABLE OF CONTENTS

PREFACE

The ability for thought is one of the most successful anthropological differences that we humans have managed to come up with, up there with language, probably. In this respect it is only natural, if a bit self-referential, for a philosophy book to be about thought and thinking. More surprising, if anything, is that this book should look at the dark side of thinking: namely that large forgotten part of our lives in which we do not think and are better off not thinking. Sometimes – actually, as this book shows, pretty often - thinking gets in the way of successfully getting things done; sometimes thinking impairs our performance, be that an action or a judgement: we often act *mindlessly* and that is also, in many cases, evolutionary advantageous for us. This phenomenon of mindlessness, I argue in this book, has long gone underestimated and while empirical psychologists have woken up to it in the last few decades, philosophers still show very little interest. Here I am not interested in a diagnosis of this bit of philosophical anthropocentrism – and indeed you won't find one in this book; more modestly, I offer a philosophical analysis of mindless behaviour which is meant to do justice to a wrongly ignored phenomenon. I present the empirical evidence for mindlessness from the last few decades of psychological experiments and argue that mindless behaviour (automatic and habitual actions, skilled activities, conventional behaviour, emotional and spontaneous reactions) is a philosophically interesting topic because it challenges some established accounts of agency, such as those based on a combination of *internalism* and *causalism* which explain human actions in terms of the rationalising psychological states that are supposed to be the causes of our actions. The positive contribution of this book is not meant to be offering a new theory of action that does justice to mindless behaviour – even though you will find, here and there, suggestions as to which direction I consider more promising; the positive contribution of this book is only meant to be to acknowledge an important part of our lives that we are all too familiar with but hardly talk about – at least in books, anyway. But acknowledging our mindlessness should not be understood to mean acknowledging our shortcomings or a less than rational aspect of our agency: this book normalises mindless behaviour in showing that it is, mostly, fully intentional and fully rational. It is not an embarrassing exception: it is the proud rule of how we cope with a

challenging world. Being mindless is just as good and central a part of ourselves as being mindful: when we do something mindlessly we are acting in just the same sense in which we are acting the rest of the time; and if the rest of the time we are acting rationally and intentionally, then we are also acting rationally and intentionally when we are acting mindlessly; if the rest of the time we are acting freely, then we are also acting freely when we are acting mindlessly. We are being ourselves and those mindless actions are also our own. Indeed, in the spirit of Aristotle, I am tempted to say – even though I shall not pursue this claim in the book – that our mindless self is our true self because it is not mediated by thought. Being mindless is, in short, a good thing. Let us not be ashamed of it; let us cultivate it.

Ulm, 13.2.2013

ACKNOWLEDGEMENTS

This book continues the work I had started while doing my PhD at the University of Edinburgh and that I published as a monograph in 2008 under the title *Mind Out of Action*. In this respect many colleagues, friends and students who contributed to that work have also been influential, more or less explicitly, for this prosecution. For that I thank them. Also, in the last year I was lucky enough to organise (together with Neil Roughley) two conferences on issues that are central to this volume: *Control and the Absent Mind* in the summer of 2012 and the *Automatic Actions* workshop in January 2013. I thank those that have made those events possible and in particular speakers and other participants at the two meetings. I was pleasantly surprised to see how many people even within philosophy are actually interested in this topic. Also, in the summer semester of 2012 I taught a course at Universität Duisburg-Essen on the philosophy and psychology of habit and automaticity which has helped me greatly in finishing this book: it was a lively and successful class and I learned a lot from my students; it was too large a class to mention all the students there by name, but at least in one case in this book I explicitly use an example that I owe to a student from that class, Gerd Szafran. I thank them all: writing books is fun but teaching is way better! I am also grateful to my university for being able to teach on topics that I am writing on: that is not a given. Finally, I must thank those who were more directly influential in producing this book: Josef Clinch for a brilliant cover; Enza Trotta for helping with the bibliography; Günter and Katharina Gerstberger for flying all the way to Mexico so that I could have the necessary quiet in their empty house to finish this book; and all those at CSP who have been very supportive: thanks go at least to Carol Koulikourdi, Andy Nercessian, Amanda Millar and Emily Surrey.

CHAPTER ONE

INTRODUCTION

Thinking is overrated. No wonder that philosophers in particular and intellectuals more in general would talk thinking up: it is both their only tool and their best product. More surprising is, if anything, that common sense should also hold thinking in such high regard: THE people should know better – experience should have taught them. *Think twice*; *think again*; *come to think of it*; *think before you speak*; *think it over*; *well thought-out*; these are only a few examples of how entrenched is thinking's reputation. And, to be sure, it would be silly to pretend that thinking is always a bad idea: sometimes you should think twice; sometimes thinking again is the better option; sometimes thinking before speaking is wise; some options should indeed be thought over and likewise some courses of actions must be thought-out. Sometimes, but not very often. Very often, in life, you know what to do already, you do not need to think: because you have done it thousand times before; because you have trained hard; because it is the right thing to do; because you have learned it through years of practice; because you are an expert, or anyway your body definitely is.

This book is about all those other times, when thinking would just be a waste of precious cognitive resources; all those other times when it's not just that you don't need to think; rather, thinking would interfere and compromise a performance which has otherwise been refined to perfection. You have probably had the experience of accidentally looking at the steps while walking down a flight of stairs: that doesn't make walking down easier; it makes it harder. The modern version of this phenomenon has to be thinking about your pin number before typing it in: that's just about the best way to get it wrong. Sometimes, as I wait at the back of the queue, I need to remind myself *not* to think about my pin number when my turn comes. The curious destiny of man: billions of years of evolution only to end up being afraid of thinking.

This book is about the philosophy and psychology of acting without thinking: the latter shows a growing body of studies on automaticity, habit, and other *mindless* phenomena; the former struggles to catch up, hanging

on to an intellectualist picture of human action. This intellectualist bias, widespread across philosophy, needs readdressing: there are growing empirical reasons to suspect that explaining action in terms of conscious (or even less than conscious) thought ends up leaving out the majority of human activity; the book presents this empirical evidence and also puts forward a series of theoretical arguments which point in the same direction. So this is, in short, a philosophy book which argues for less thinking: a bit like a tobacco corporation lobbying for fewer cigarettes.

There is also another important sense in which philosophy overestimates thinking: normatively. There is a long tradition in philosophy, often referred to as Kantian, which takes moral agency to require rational deliberation; or, to formulate the same kind of view more carefully, which at least prioritises, morally speaking, actions which are the result of rational deliberation over those which are not. This diverse tradition has possibly its most explicit manifestation in the so-called *Doctrine of Double Effect*, according to which whether or not an action is intended makes a difference to its moral permissibility.[1]

In fairness, at the normative level we do not find the kind of intellectualist monopoly that we just described for the theoretical level of action explanation. First of all, this is because of the wide-spread influence of consequentialism. But also, and this is more to the point here, because often the intellectualist Kantian approach is contrasted with an Aristotelian approach founded around character, virtue, and habit. According to a particular interpretation of Aristotle that I am very fond of, a certain degree of habit, automaticity, or *mindlessness* is even necessary for an action to count as virtuous: quickly said, the idea would be that for an action to count as virtuous it is not enough that the action be in accordance with the virtues or what the virtuous agent would have done in that same situation; rather, the action is only virtuous if it is a spontaneous expression of the agent's character; if the agent did not need to think about whether to act that way because the virtuous action came naturally (automatically, habitually, or mindlessly) to her. So at least at the normative level, and if one follows a particular streak in the Aristotelian tradition, thinking has not always been overrated in the history of philosophy.

It is important to emphasize from the start the clear parallel between the theoretical overestimation of thinking and its normative overestimation: on the one hand, an appeal to conscious (or less than conscious) thought is

[1] More on double effect in Chapter 7: also, I am writing a book on double effect which I hope to bring out next year and which is provisionally titled *Ethics without Intention*.

taken to be necessary for the *explanation* of human action; symmetrically, an appeal to thought is taken to be necessary for the *justification* of human action. That this issue transcends the traditional divide between theoretical and practical philosophy could well be an indication of its centrality. But in this book I will, with the only exception of Chapter 7, merely deal with the theoretical role of thinking, across philosophy and psychology, in the explanation of action.

Here a first distinction is in order, which also reflects the structure of this volume: I would argue that philosophers overestimate thinking and that they therefore underestimate the diverse phenomenon that gives this book its title, mindlessness; the same cannot be said of psychologists, some of which at least in one important respect actually overestimate mindlessness and automaticity: that is because they take these phenomena to be challenges to free will. So in Part I and Part II of this book I discuss mindlessness and automaticity from within a philosophical framework, arguing that standard philosophical theories for the explanation of human action have a problem in dealing with mindlessness and automaticity; while Part III of the book is dedicated to the psychologists' overestimation of mindlessness and automaticity, where I argue that these phenomena pose no interesting challenge to free will. The continuity between Part I & II and Part III of this book should be understood as a general attempt at normalizing mindlessness: on the one hand, mindlessness needs to be philosophically normalized to show that it is a common phenomenon which cannot be banned as anything less than fully intentional and fully rational; so we need philosophical theories of action that can account for mindless behaviour as fully intentional and fully rational behaviour. On the other hand, mindlessness needs to be normalized also within empirical psychology in that it is, again, a common phenomenon which has nothing to do with classic challenges to our free will: briefly, the claim is that when we act or judge mindlessly, we are no less free than when we act or judge mindfully.

In the rest of this introductory chapter I present the topic of this book, mindlessness, in its three crucial elements: its functional aspect (Section 1); its explanatory aspect (Section 2); and its normative aspect (Section 3). I then conclude this introduction by briefly summarizing the content of each chapter (Section 4).

1. Mindlessness: functional

Following up on the framework of the philosophical underestimation of mindlessness with which we have started this introduction, we could

very roughly say that while the normative aspect of mindlessness has a solid old Aristotelian tradition and has also experienced a recent revival in virtue ethics, the explanatory and functional aspects of mindlessness remain under-discussed in philosophical circles. That is all the more surprising when one considers the amount of empirical work that has gone in the psychology of mindlessness in the last decades, including some recent successful popularizations of this work. Starting from Kahneman's *Thinking, Fast and Slow* through Gigerenzer's *Gut Feelings* to Duhigg's *The Power of Habit*, the general public shows more interest for this topic than philosophers. Those books are only meant to offer a feel for recent interest in the topic and should not be considered a coherent group (as for example one could easily add to the list at least Gladwell's *Blink* and also *Nudge* by Thaler and Sunstein), neither in terms of content nor in terms of target: while the first two of them – along with *Nudge* - have been written by established psychologists, *The Power of Habit* – just like *Blink* - is just a journalist's journey in both empirical research and common sense on the role of habits and habitual behaviour in our life.

Let us start with the functional aspect of mindlessness. Here is an astonishing example of the kind of research that has lead to this small storm of popular interest: Sian Beilock and colleagues compared the performances of expert golfers with those of novices and found that, under time pressure, expert golfers perform better than when they are not under time pressure! Novices, unsurprisingly, perform better when they are not under pressure. In a related study, expert golfers and novices were either told to concentrate on their swing or they were given an extra task to concentrate on (counting the number of tones coming out of a recorder). Expert golfers did worse when they were able to concentrate! Unsurprisingly, novices did better when they could focus solely on their swing.[2]

This study is important for two related reasons. First of all, this study poses a challenge to common sense: thinking before doing is not always the best strategy. It's not just that thinking before doing isn't always necessary and therefore we should sometimes abstain for reasons of economy: rather, sometimes thinking before doing makes our performance worse. Thinking, then, is not always functional to doing; rather, as Beilock shows, thinking is sometimes dysfunctional to doing. But for the purposes of this book the criteria for the functionality and dysfunctionality of thinking before doing are even more interesting – and that's the second reason why this study is important: whether or not thinking is functional or

[2] For more details on this study see the next chapter.

dysfunctional depends, namely, on who you are; what you have learned; what you are good at; what your habits are like. Importantly, then, whether or not thinking is a good idea does not depend on what you want or on your psychological and conscious states at the time of the performance: it rather depends on your history. If you have been educated in a certain way, for example, thinking will not be a good idea – it will be a bad idea (in this case, if you have been educated to the game of golf). Notice, also, that Beilock's study confirms the experiences I mentioned at the very beginning: don't think nor look at your steps while walking on a familiar path or flight of stairs, otherwise you will increase the chances of tripping over. It's not just that it is a waste of time and cognitive resources; it is positively dangerous, as Beilock and colleagues showed with the example of expert golfers. That's because most of us, most of the time, in most places, are expert walkers.

We have thereby identified two important elements of mindless performances: mindless performances are often either skilled activities or habitual activities – both of which have to do with expertise. Here I do not want to attempt a definition of mindless action, and certainly I do not think that mindless action must necessarily be either skilled or habitual; but it does seem to me that most of its more interesting manifestations have to do with expertise, without ruling out other forms of unconscious or automatic behaviour, such as spontaneous reactions or emotional behaviour. A further interesting related kind of activity are actions that follow social conventions – think of Hart's famous example (in philosophy anyway) of taking the hat off when entering church; while those behaviours need not be unaware, it seems that they are often mindless in that what explains these behaviours is the social convention itself and not the psychological states of individual agents. I do not want to mention too many examples in this introductory chapter because this book is full of either real or imaginary examples of mindless behaviour, automatic actions, habits, and related phenomena. But in thinking about skilled activity and habitual action, one could think on the one hand about the movements and strokes of a tennis player and, on the other hand, about one's morning routines. It is quickly clear that skills and habits are related phenomena: both are the result of endless repetition, refinement, improvement.

A similar point about expert intuition in everyday life can be found in Kahneman:

> Expert intuition strikes us as magical, but it is not. Indeed, each of us performs feats of intuitive expertise many times each day. Most of us are pitch-perfect in detecting anger in the first word of a telephone call,

> recognize as we enter a room that we were the subject of the conversation, and quickly react to subtle signs that the driver of the car in the next lane is dangerous. Our everyday intuitive abilities are no less marvellous than the striking insights of an experienced firefighter or physician – only more common (Kahneman 2011:11).

Interestingly, Kahneman's examples are not only cases of physical action such as in Beilock's golfers; there are also, importantly, cases of expert judgement. One famous case in point, here, also mentioned by Kahneman, is Klein's (1999) fire-fighter who, without knowing why, gets all his men out of a burning building the moment before the floor collapses. The fire-fighter chief realised only after the event that the fire had been particularly quiet and that his ears had been particularly hot. The chief had picked up those features of the situation without realising it. That's the expert feat of someone who has been in thousand of burning buildings before; someone who has heard and smelled thousands of different fires before. The case is very similar to Beilock's golfers once you think of it: had the chief stopped to reflect on exactly what was different, he might have delayed his decision for those very few seconds that saved his life and the lives of his men. Again, for the fire-fighter chief as for the expert golfers the same is true: thinking would not have just been a waste of cognitive resources that one could have better invested somewhere else; thinking would have been fatal to the successful completion of the task.

Sometimes this fire-fighter story gets told somewhat differently. In this variant, one could imagine that the same fire-fighter makes the same correct call, rescuing his crew. But afterwards, the fire-fighter is still at a loss as to what made him get his crew out of the building. That is, the fire-fighter himself cannot tell, but someone else – maybe a fire-fighters' trainer or some other sort of meta-expert, maybe some cool CSI-type from the telly – is able to explain to the chief that he must have noticed – at some level – that the fire was particularly quiet and that his ears were particularly hot. And here one could introduce a further variant: where in one version of the story the chief is able to confirm the meta-expert's hypothesis while in the other version of the story the chief cannot tell either way whether the meta-expert has correctly identified the actual features of the situation that made his call the right call.

Why am I telling all these different hypothetical stories about the fire-fighter chief? While which one of these different stories one takes does not make a difference to the functional aspect of the chief's correct mindless call, those different stories can be distinguished philosophically. In the last one, for example, one would have to at least challenge the idea that the

chief deserves praise: he has no idea why he made the call and has no privileged access to his reasons. Why should he then be responsible for the call? Furthermore, one could imagine that in the last version of the story it is whoever trained the chief that deserves the credit and not the chief himself, who blindly trusted his training. On the other hand, though, if we think of a reversed case where the chief acts similarly mindlessly but makes the wrong call, would the mindlessness of his action excuse him? Probably not; but, apart from notorious asymmetries between positively and negatively evaluable behaviours (the so-called Knobe Effect), the question would remain of whether one could truly claim that the chief intentionally and rationally made the call (right or wrong) when he had no access to his reasons for making that call. This introductory chapter is not the place to get into the argumentative details of these issues; here I just wanted to illustrate the point that right behind a quite obvious functional role of expertise there are difficult philosophical issues to settle (more on this also in the next section) – and that's what this book attempts to do.

There is another important element to the functionality of mindless actions and mindless judgment – which I think can also be nicely illustrated by the fire-fighter's example. One crucial issue is whether deliberation is necessary for rationality and, therefore, whether mindless actions and mindless decisions can be rational ones. And in one sense I have already answered this question by illustrating the functionality of mindlessness. Did the fire-fighter chief act rationally? He certainly did, one could answer, in that he did the right thing and saved his crew. But what if things had turned out differently? It can't be that the rationality of his behaviour depends entirely on how things turned out, it will be quite fairly pointed out. And then one would have to start assessing things such as the reliability of the fire-fighter's expertise in order to establish whether the fire-fighter acted rationally. But in one important sense this is beside the point: as the fire-fighter acted without knowing what his reasons for acting were, one may think that that's enough to preclude this action from being rational. Mindless action, then, also throws up important questions about the nature of rational agency and rational decision. And it looks like one will need a conception of rationality that allows for at least some cases of mindless action and mindless decision to be included. Here, again, an intellectualistic conception of rationality risks leaving out a lot of successful human activity. And, I'd like to point out to conclude this section, an intellectualistic conception could also leave out some unsuccessful human activity.

Here is a scenario which I think illustrates this point quite nicely and that I owe to a student of mine: some people – especially older people –

tend to take their umbrella even if it's not raining. That, you could say, is a good example of a habit: as you walk out of the door, you mindlessly grab the umbrella. Now, you may also think that this scenario illustrates how habits can lead us to make mistakes and how sometimes we consciously need to intervene to prevent our habits from leading us astray. This is, indeed, what Kahneman suggests with his talk of System 1 (the automatic level) and System 2 (the conscious level): the conscious level has a monitoring role over the automatic level and sometimes needs to intervene when the automatic level cannot cope alone. Now, if you saw an older gentleman carrying an umbrella on a mild day, you may find yourself wondering whether his conscious System 2 should not have known better. What a waste of time and energy, taking your umbrella on a mild day. But the point of this example is to show that this way of thinking is short-sighted: one should not think about the negligible amount of energy being wasted on this one day; one should think long term about the substantial amount of energy saved by the establishment of the habit: instead of checking, everyday, what the weather's like; instead of watching, everyday, the weather report; and instead of all the risks connected with these activities (the report might be inaccurate, say; the weather may unexpectedly take a bad turn); instead of all of this, there is the more economical and safer option of establishing a practice. This is only a tiny example of what I take to be the fundamental role of habits, practices, and automaticity (in one word, mindlessness) in successful economical behaviour and life in general: that's why just deeming this behaviour as something less than rational won't do.[3]

Summing up, the point about the functional aspect of mindlessness is that, whatever your teacher used to tell you at school, thinking before doing, thinking before deciding and thinking before speaking are not always your best bet.

2. Mindlessness: explanatory

In talking about mindless phenomena in terms of their functional, explanatory, and normative roles, one could be easily tempted to group the explanatory and functional roles together as opposed to the normative role. But that would be a bit quick; and indeed a different grouping is, in one important respect, more accurate: namely explanatory on the one side and

[3] Just to be absolutely clear: this is clearly not to say that there are no bad habits and neither is to say that we should never consciously intervene to stop or correct a mindless activity of ours.

normative and functional on the other side. Because, quite obviously, what I said in the previous section about the functionality and dysfunctionality of thinking was normatively loaded. To say that an agent is better off not thinking is a normative statement; on the other hand, though, to say that a performance is more or less successful as a result of deliberation or lack thereof doesn't need to get one entangled in complicated normative issues: the success of a performance can just be measured in terms of what the agent was trying to do, so that both expert golfers and novices were trying to get a good hit, and in these terms we can say that while novices are better off concentrating and thinking about the performance and taking their time, expert golfers are better off not concentrating about what they are trying to do.

The explanatory aspect of mindlessness and automaticity, on the other hand, does not involve any of these evaluative considerations. Here the point is, quite simply, how one is to explain these agential phenomena. And, philosophically, the first issue is exactly whether those phenomena should count as actions or not. Briefly stating a possible objection, if action necessarily requires awareness and we define mindless behaviours in terms of lack of awareness, then mindless behaviours cannot be actions. This conceptual point, though, appears to fly in the face of the kind of cases that we have been introducing. Quite obviously, as we go through our morning routine of turning the alarm off, brushing teeth, putting the kettle on, turning on the radio, and so on, we are acting. And we are acting in just the same sense in which we are acting the rest of the time. The same goes for the expert golfers, the tennis players, the fire-fighters, etc. Intuitively, those are actions. But then we need an account of human action that can include all these mindless activities.

Most cases of mindless behaviour are very obviously cases of action, as we just saw. But it's not always that easy. Think of Damasio's Iowa Gambling Task, for example. Participants were confronted with four different decks of cards to which positive and negative rewards were attached. By playing the game and turning enough cards over, healthy participants realized pretty quickly which decks were advantageous and which less so. The amazing thing with this study, though, is that participants had already altered their behaviour to their own advantage before being able to say which the better decks were. And, even more astonishingly, participants' hands seem to have noticed the difference between the decks even before that. Damasio and colleagues therefore identified three different phases: a first phase - "pre-hunch" - where participants had yet to behave advantageously but were already showing skin conductance responses (which is a measure of sweat and arousal)

before turning the risky cards; a second phase - "hunch" - where participants couldn't yet verbally say which decks were more advantageous but were already acting advantageously. And a third "conceptual" phase where subjects were finally able to tell the experimenter what was going on.

This study raises all sorts of questions, but the most obvious one for our present purposes seems to be, again, that this is a case of functional behaviour which cannot be obviously explained in terms of conscious deliberation. Here it is worth introducing another distinction that will be helpful throughout the book: in many of the intuitively easy cases of expertise, habit, and automaticity that we have already mentioned, whether or not the agent is deliberating about her performance and whether or not the agent is presently conscious of her activity, the agent has at least direct access to her performance: she can retrieve the relevant information at will; she can focus or refocus her attention to the performance at will; she needs no external help[4]. Sometimes philosophers talk, in these cases, of dispositional awareness or access consciousness, to distinguish these cases from cases where an agent is wholly unaware of an aspect of her performance (more on this in the next chapter so don't worry if this goes a bit too quick here).

To illustrate genuine unawareness, take the following standard case (from Donald Davidson) which will feature again in more detail in later chapters: I board a plane marked as headed to 'London' with the intention of flying to London, England. Unbeknownst to me, the plane is actually headed to London, Ontario. I am unaware of a relevant feature of my action, namely that the plane I am boarding is actually headed to London, Ontario. And I have no direct access to that bit of information: I cannot retrieve it in my memory and it is also, we can imagine, not before my very eyes at this moment. This would be a case of genuine unawareness as opposed to being a case of dispositional awareness or access consciousness – as for example in the case in which I mindlessly follow the queue for the London, England plane. I am not thinking all the time about the fact that the plane I am about to board is headed to London, England; nor am I looking all the time at the screens displaying the information that the plane is in fact headed to London, England. But even though I am not occurrently aware of this fact at this very moment because I am not thinking about it, I am at least dispositionally aware of it (or access conscious of it) and that explains my mindless behaviour. So there seems to be an obvious difference between the case in which I notice that

[4] Here there would be interesting and important things to be said about the difference that the extended mind hypothesis makes to this kind of distinction but I will have to leave this one issue to the side here.

the plane is indeed headed to London, England at the one end of the spectrum and the case in which I, unbeknownst to me, board the plane to London, Ontario while thinking it is headed to London, England at the other hand of the spectrum. But, as Damasio's study shows, there is probably quite a lot of interesting and difficult agency between those two extremes; and at least some of this agency will be mindless.[5] And trying to establish what we should say about those sorts of cases is one of the main tasks of this book.

As the main topic of the book is exactly the explanatory aspect of mindless behaviour, here I don't want to say too much about this issue, rather just give a feel for its importance: the idea that at least parts of human agency are not necessarily conscious is a fundamental anthropological challenge on top of a challenge to philosophical theories of action. One important issue, for example, and one that this book does not discuss, is the relation between mindless human agency and animal actions. Whether or not animals can be said to truly act is an open question, but if one gives up – as this book suggests – on the thinking and conscious requirements on human agency, then this step could also be interpreted as one in the direction of animal agency. And indeed non-human animals are in one respect also experts: they can learn by repetition and improve performance. Is there a form of continuity between animal action and automatic (human) action? Here it is interesting to note the asymmetry between the genotypical and phenotypical levels: one could imagine that the natural history of behaviour is one where thinking before doing and conscious behaviour come, historically, after doing without thinking and unconscious behaviour; on the other hand at the individual level we observe the opposite: through practice and habit formation, agents start having to think less and less. A performance that used to require thinking and consciousness can then later be successfully completed mindlessly.

Let me mention another huge issue that is closely related to the topic of this book but that I again won't have the occasion to go into in any detail: there is a tradition, both in philosophy and in empirical psychology, of distinguishing between so-called rationalism and so-called sentimentalism; this tradition goes beyond the divide between theoretical and practical philosophy too in that both in talking about the justification of action and talking about the explanation of action people tend to divide themselves between rationalists and sentimentalists – traditionally, say, Kant would be

[5] Other cases that are in between the two clear extremes are the many experimental cases of automaticity to be found in the priming literature, many of which are discussed in the next chapter.

a prominent rationalist while Hume would be a prominent sentimentalist, just to drop some huge names there – in empirical psychology, think of the influential research conducted by Green or by Haidt on this traditional divide at the explanatory level. This issue is obviously related to the distinction between mindfulness and mindlessness that is at the centre of this book but I do not think that it should be confused with it: mindlessness, to be absolutely clear, does not imply a form of sentimentalism – while on the other hand the kind of intellectualism that I criticize in this book is often to be found in rationalists and I have myself mentioned Kant as opposed to Aristotle in this very introductory chapter.

So let me say, prudently, that the arguments in this book are supposed to be silent on the divide between rationalism and sentimentalism; which I would distinguish from the divide topical to this book between what one could call *intellectualism* and, for lack of a better word and out of respect towards our founding father, *Aristotelianism*.

3. Mindlessness: normative

It is natural to think that there is a close relation between agency and morality: for example, it is often suggested that someone can be praised or blamed only if she was conscious or aware of what she was doing. Diminished states of consciousness or lack of intention are, in many legislatures, either grounds for justification, excuse, or at least diminished sentences. As the *mens rea* principle suggests, the psychological state of the agent is a crucial element in the evaluation of her actions. Without getting into issues in legal philosophy, this is only meant to point out a first important issue in the normative evaluation of the phenomenon of mindlessness, and one that is also closely related to the point in the last section about whether mindless behaviours should count as proper full-blown actions.

This book, though, leaves all these issues of responsibility for mindless behaviour and evaluation of mindless performances to the side, concentrating on the theoretical problems to do with the explanation of mindless performances. That those issues cannot be covered in this book doesn't mean, though, that the author considers them any less important: think, for example, of the fascinating problem of when we should start praising or blaming the agent in the Iowa Gambling Test. Already by the time her hands start sweating, even though both her behaviour and her verbal reports indicate that she doesn't consciously realise that the decks are different? And would it be fair to blame an agent who alters her behaviour to her own advantage later than other agents, if none of these

agents can report, by that stage, verbally on the structure of the game and the quality of the decks?

And what about good and bad habits? Should the fact that a performance has become habitual constitute some sort of excuse for that behaviour? Imagine you do actually meet the old gentleman of the story with his umbrella on a mild day and you challenge him as to why he's carrying an umbrella on a mild day. Is his explanation in terms of habits and established practices going to count as a justification or at least an excuse? Carrying an umbrella is a pretty innocuous activity (normally, anyway), but you can easily think about habits that are not that innocuous. Imagine that a guest who you know to habitually speak very loudly wakes up your son sleeping upstairs. Should you show for this guest more understanding than for someone who normally speaks very softly but today is surprisingly loud and also wakes up your son?

Those are just stories, but at least the following happens to a lot of people and has maybe already happened to you too: you are driving on your usual route to work. This time, though, you are taking your partner to the airport and it just happens that in order to drive to the airport you need to drive part of your usual route to work. As you are driving on the familiar part of the route, your performance mindlessly adapts to your habitual practice of driving to work and you take the wrong turn: you haven't forgotten that you are driving to the airport; it's just that your habit was stronger and made you take the normal turn on your route to work. You mindlessly slipped into your habit: the mechanism, we can imagine, is similar to the umbrella case. It just happens that, statistically, the overwhelming majority of the time, when driving on this road, you are driving to work; that's why the kicking in of the habit is statistically efficacious but leads you astray this one time. But now the mechanism isn't really what interests us: rather, what should we say, normatively, about this case? Suppose your partner misses her flight because you took the wrong turn. Can you justify yourself or at least excuse yourself by explaining what happened and talking about acting automatically or out of habit? Is that a legitimate excuse or justification?

Here there are various possibilities that I cannot explore in any detail (a bit more on this in the concluding chapter though), but one could suggest at least the following:

1) Your mindless driving is a full-blown action with all its consequences – no excuse, even though you acted unintentionally or at least without the relevant intention and therefore your partner

should not assume any bad will on your part (more on this in the discussion of the role of intention in ethics in Chapter 7);

2) You didn't really mean to take the turn, your habit overpowered you. That's no justification but it is an excusing consideration: your partner should show some understanding;

3) The story about the economic advantage in the long term that we told about the umbrella applies here too: you are responsible for the establishment of a practice that is advantageous to you; you knew or at least should have known the negative consequences of the practice, which are, though, outweighed – at least objectively but probably also from your own subjective point of view - by its positive consequences: here the crucial counterfactual question may have to do with whether you would have been prepared to renounce the whole practice only to avoid your partner missing her flight this one time. And so even if now you acted mindlessly, that is really part of your overall plan to act economically and therefore you are to blame for the foreseeable consequences of your actions.

Those and similar considerations are clearly relevant to the evaluation of the phenomenon that is at the centre of this book; but I will leave these issues for future discussion and will not analyse them further in this volume (with two exceptions: Chapter 7 is dedicated to the role of intention in ethics; and the concluding Chapter 10 briefly takes up some of these questions about responsibility for mindless behaviour).[6]

4. Summary of the book

Let me briefly say what is to come in this volume: the book is divided into three parts: Part I looks at mindlessness from both a philosophical and psychological point of view. I present some of the more influential empirical evidence and discuss its philosophical relevance. In particular, in Chapter 2 I look at empirical work on habits and priming and ask what kind of consequences these empirical findings have for philosophical theories of action. In Chapter 3 I continue the discussion of Chapter 2, bringing the focus away from the psychological lab and into the real world of our daily lives, which are full of habits and automatic actions. Can the

[6] There is little ethics in this book, but readers who should be interested in my writings in ethics and applied ethics may look at the following: Di Nucci 2009b and 2009c; Di Nucci 2011c; Di Nucci 2012b; Di Nucci (forthcoming a-f); and Di Nucci (book manuscript).

influential causal theory of action put forward originally by Davidson and dominant in contemporary philosophy account for habitual and automatic actions? In Chapter 4 I look at the causal theory of action from a different point of view which complements my argument against it.

In Part II of this volume I discuss the consequences of mindless action for the idea that intention and action are closely related to each other: can there be action without intention? If intentions were necessary for action and mindless behaviour is not intended, then that's a big challenge to the normalization of mindless behaviour that motivates this book. That's why in Chapter 5 and Chapter 6 I discuss, in particular, the so-called Simple View of intentional action, according to which an intentional action is necessarily intended: I argue against this view. Then in Chapter 7 I look at the normative side of the relation between action and intention and in particular at the idea that whether or not an action was intended makes a difference to the moral permissibility of that action.

Finally, in Part III, I look at the consequences of the empirical evidence I present and the theoretical arguments I put forward for the free will debate: in particular, in Chapter 8 I look at the relationship between the causal theory of action and a compatibilist approach to free will, arguing that giving up on the former does not imply having to give up the latter. Then in Chapter 9 I argue that empirical research on mindlessness presents no challenge to our free will by looking in particular at the example of research on priming. Finally, in the concluding Chapter 10 I briefly discuss some of the normative issues that emerge from this book and that I consider worthy of future work.

PART I

AUTOMATICITY

CHAPTER TWO

HABITS, PRIMING, ALIEFS AND THE EXPLANATION OF ACTION

There is a growing body of evidence on the influences of automatic and unconscious processes on our actions. Here I introduce some representative examples of this growing body of evidence, chosen so as to form a diverse group of related *mindless* phenomena: habits, skills, priming and nudges. I then argue that this evidence challenges traditional belief-desire-based approaches in the philosophy of action. I further discuss a recently proposed solution to this challenge, Gendler's *Alief*, finding it wanting. I conclude by sketching my own alternative solution, based on the old story of Buridan's ass.

1. The empirical case for automaticity

In a fascinating study on eating habits, consumption of fresh and stale popcorn at the cinema by habitual cinema-going popcorn consumers was tested: habitual popcorn consumers ate just as much one-week-old stale popcorn as fresh popcorn (Neal, Wood et. al. 2011). Those without the habit of eating popcorn at the movies ate more fresh popcorn than stale popcorn. Neal, Wood et al. refer to this sort of habitual behaviour as 'automatic', 'non-goal-dependent', and say that it is not under 'intentional control' (which once they refer to as "personal control" (2011: 9)).

Neal, Wood et. al. (2011) found that either an unusual environmental context (eating popcorns in a meeting room instead of a cinema) or a novel way of carrying out the habit (eating with the non-dominant hand) disrupted habitual performance, resulting in even the subjects with the cinema-going popcorn-eating habit eating more fresh than stale popcorn. This is taken to confirm their hypothesis that "habits should not be *activated* automatically outside of their typical performance context and should not be *executed* automatically when responses are performed in novel ways" (2011: 1). This suggests, then, that habits are normally both activated and executed *automatically*. What does it mean that habits are activated and executed automatically? Here is again Wendy Wood: "Cue-

response associations are basic to habit performance. By these cognitive associations, habits can be performed automatically, with little thought or effort" (2012: 980).

Habits are not the only kind of automatic performance. In walking down a flight of stairs, you are ill advised to look at your steps: should you do that, you will likely trip. When taking cash from an ATM, you are ill advised to think about your pin, just type it in. These are not the results of experiments in behavioural psychology, rather just some of those things an average person learns in walking this earth; but these could as well have been experimental results: expert golfers are similarly ill advised to take too long to put or concentrate on their swing.

> In an experiment, novices and expert golfers were studied under two conditions: they had either only up to three seconds for each putt or all the time they wanted. Under time pressure, as mentioned, novices performed worse and had fewer target hits. Yet surprisingly, experts hit the target more often when they had less time than when they had no time limit. In a second experiment, players were either instructed to pay attention to their swing or distracted by an unrelated, second task (counting tape-recorded tones). When they were asked to pay attention to their swing, as one might expect, novices did better than when they were distracted. Yet with experts, it was again the opposite. When experts concentrated on their swings, their performance decreased; when experts' attention was distracted, their performance actually improved (Gigerenzer 2007: 33; for details on the two experiments see Beilock et al. 2002 & Beilock et al. 2004).

These experiments suggest that more time to think, more attention, and more concentration worsen the performance of experts while they improve the performance of novices. Given the same type of performance, its tokens are importantly different when performed by experts and novices; and this difference seems to depend on the relevant skills that the experts have acquired and that the novices (still) lack. Skills, then, are peculiar in that not only does their exercise run to successful completion without needing thought, attention, or concentration; but the successful completion of an exercise of skill would be hindered by thought, attention, or concentration.

When it comes to our habits and daily routines, we are the experts. So that the average person is, normally, an expert at climbing stairs and typing her own pin number just like the professional golfer is an expert at putting. The connection between skills on the one hand and habits and routines on the other isn't mysterious: both skills and habits develop through time by practice and repetition, which suggests that mastering a

practice has to do with freeing cognitive resources which are no longer necessary (and which can become counterproductive, as we have just seen).

> Habits develop when people give a response repeatedly in a particular context and thereby form associations in memory between the response and recurring context cues (Neal, Wood et al. 2011: 1).

Priming offers us a further example of automatic behaviour. One of the classic examples is walking slower (than a control group) out of the experiment's room after having taken a linguistic test containing a disproportionate amount of words related to the 'elderly' stereotype: "worried, Florida, old, lonely, grey, selfishly, careful, sentimental, wise, stubborn, courteous, bingo, withdraw, forgetful, retired, wrinkle, rigid, traditional, bitter, obedient, conservative, knits, dependent, ancient, helpless, gullible, cautious, and alone" (Bargh et al. 1996: 236).

A very similar one, presented in the same seminal paper (Bargh et al. 1996), has subjects primed with a 'rudeness' stereotype who go on to interrupt the experimenter more often than two other groups of subjects (one primed with a 'politeness' stereotype and a control group). As I discuss this case in some detail throughout, I think it is important to be specific about the content of this experiment. For the rudeness priming the following words were used: "*aggressively, bold, rude, bother, disturb, intrude, annoyingly, interrupt, audaciously, brazen, impolitely, infringe, obnoxious, aggravating,* and *bluntly*" (234). For the politeness priming the following words were used: "*respect, honor, considerate, appreciate, patiently, cordially, yield, polite, cautiously, courteous, graciously, sensitively, discreetly, behaved,* and *unobtrusively*" (234). For the control group the following words were used: "*exercising, flawlessly; occasionally rapidly, gleefully, practiced, optimistically, successfully, normally, send, watches, encourages, gives, clears,* and *prepares*" (234).[1]

[1] A lot of the priming experiments have not passed replications over the last few years, and there is growing talk of the whole priming literature being discredited. Having said that, there have also been recent successful replications. I don't want to minimize this issue, which is obviously very important and all the more so for a book like mine which uses a lot of evidence and examples from the priming literature. But as far as I can see there is no consensus that priming experiments should no longer be taken seriously and I am myself not really in a position to judge on that. For those interested in this issue, there is a very informative recent article in The Chronicle of Higher Education, with references to all the relevant literature included successful and unsuccessful replications: http://chronicle.com/article/Power-of-Suggestion/136907/.

There are many other similar cases: "showing suitably primed subjects a picture of a library leads them to speak in quieter tones; showing them an image of an elegant dining room—or exposing them to the smell of soap—leads them to eat more neatly. Subliminal visual priming with an image of an African-featured face leads subjects to respond more aggressively to certain sorts of provocation. Priming subjects with thoughts of their (achievement-oriented) mother leads them to persist longer at word-find tasks; priming them with thoughts of a friend makes them more likely to help a stranger" (Szabo Gendler 2008: 659-660).

The elderly and rudeness stereotypes type of priming belong to what is now in the psychological literature commonly referred to as so-called 'concept priming', as opposed to 'goal priming'. An example of the latter is provided by an experiment where participants are primed with a high-performance goal by having them exposed to the following words: "*win, compete, succeed, strive, attain, achieve,* and *master*. In the neutral priming condition, these words were *ranch, carpet, river, shampoo, robin, hat,* and *window*" (Bargh, Gollwitzer et al. 2001: 1016). Those primed with a high-performance goal did then much better than the control group on a subsequent intellectual task.

Something similar happens with so-called *nudges*:

> Carolyn is the director of food services for a large city school system... One evening, over a good bottle of wine, she and her friend Adam, a statistically oriented management consultant who has worked with supermarket chains, hatched an interesting idea. Without changing any menus, they would run some experiments in her schools to determine whether the way the food is displayed and arranged might influence the choices kids make. Carolyn gave the directors of dozens of school cafeterias specific instructions on how to display the food choices. In some schools the desserts were placed first, in others last, in still others in a separate line. The location of various food items was varied from one school to another. In some schools the French fries, but in others the carrot sticks, were at eye level. From his experience in designing supermarket floor plans, Adam suspected that the results would be dramatic. He was right. Simply by rearranging the cafeteria, Carolyn was able to increase or decrease the consumption of many food items by as much as 25 percent. (Thaler & Sunstein 2008: 1)

Priming effects could be considered the psychological correlate of nudges: they alter behaviour by altering an individual's psychology, rather than the environment. The *choice architecture* (that's how Thaler and Sunstein talk about nudges) of priming is more fundamental, it could be said, because it shapes the *agent* of choice rather than the *object* of choice.

On the other hand, just like with the habits experiments that we have just looked at, nudges appear to affect behaviour by intervening on the environmental cues that trigger the behavioural responses in automatic and habitual actions.[2]

The idea behind nudges is for political institutions to appropriate old and well-known marketing tools. The following experiment may provide a good illustration of how the private sector has been using 'nudges' all along: Iyengar & Lepper (2000) presented shoppers with two tasting booths with either six or 24 different varieties of jam. While more shoppers stopped at the booth with more varieties of jam, more shoppers bought from the booth with less varieties of jam (see also Gigerenzer 2007: 31). This experiment too appears to confirm an often heard commonplace according to which too many alternatives make choosing difficult.

In the next section I introduce traditional approaches in the philosophy of action that, I argue in the section after next, are challenged by the empirical evidence on automaticity just presented.

2. Traditional action theory

To avoid unnecessarily overcomplicating my discussion, here I will focus throughout only on the experiments on priming and habits, and in particular on the 'interrupting' case and the 'popcorn' case.

To interrupt someone else is something that one may do both intentionally and unintentionally. Take the following case: I know that my partner is rehearsing at home with her colleagues for an important concert and I know that they don't like to be interrupted. Still, I have left some important documents on my desk at home which I will need later in the day, so I decide to go back home to grab these documents even though I know that I will thereby interrupt their rehearsal. I take it that this is an example of intentionally interrupting someone else. Compare this with a case in which I come back home without knowing that they are rehearsing (it's not as if I had forgotten, they never told me): again, I interrupt them by coming in. This second case, I take it, should count as unintentionally interrupting someone else.

How does traditional action theory account for the first case being intentional and the second case being unintentional? Let us look at a traditional account due to Donald Davidson. On Davidson's account some

[2] I have recently written on nudges and mindlessness elsewhere: see Di Nucci (forthcoming g).

action A is intentional under description *d* only if that action was caused by a primary reason of the agent comprising of a pro attitude towards actions with a certain property, and a belief that action A, under description *d*, has that property[3]:

> R is a primary reason why an agent performed the action A, under description d, only if R consists of a pro-attitude of the agent towards actions with a certain property, and a belief of the agent that A, under the description d, has that property. (Davidson 1980: 5)

What Davidson says about 'primary reasons' helps us understand their role for intentional action: "Such a reason gives minimal information: it implies that the action was intentional" (1980: 6); also: "To know a primary reason why someone acted as he did is to know an intention with which the action was done" (1980: 7).

The relation between Davidson's 'primary reasons' and *intentions* shows that Davidson's account is the reductive counterpart of the so-called *Simple View* of intentional action: E φ-s intentionally only if E intended to φ: 'For me intentionally to A I must intend to A . . . I will call this the Simple View' (Bratman 1987: 112; on the Simple View see also McCann 1991, 2010, 2011 and Di Nucci 2009a & 2010a – Chapter 5 and Chapter 6 are dedicated to this and related issues).[4]

---

[3] Just to note that Davidson offers only necessary conditions to avoid the problem of deviant causal chains. Much more on this in Chapter 4.

[4] On Bratman's own Single Phenomenon View, there will be an intention in the causal history of all intentional actions. And, furthermore, an action *A* can be intentional only if it is included in the 'motivational potential' of some intention, even if the intention does not necessarily need to be an intention to *A* (Bratman 1987, 119–120). For a possible way of filling in a view of this kind, see Mele & Moser: "Necessarily, an agent, S, intentionally performs an action, A, at a time, t, if and only if: (i) at t, S A-s and her A-ing is an action; (ii) at t, S suitably follows-hence, is suitably guided by-an intention-embedded plan, P, of hers in A-ing; (iii) (a) at the time of S's actual involvement in A-ing at t, the process indicated with significantly preponderant probability by S's on balance evidence at t as being at least partly constitutive of her A-ing at t does not diverge significantly from the process that is in fact constitutive of her A-ing at t; or (b) S's A-ing at t manifests a suitably reliable skill of S's in A-ing in the way S A-s at t; and (iv) the route to A-ing that S follows in executing her action plan, P, at t is, under S's current circumstances, a suitably predictively reliable means of S's A-ing at t, and the predictive reliability of that means depends appropriately on S's having suitably reliable control over whether, given that she acts with A-ing as a goal, she succeeds in A-ing at t." (1994: 63). I do find this account of intentional action ambiguous as to its relationship to the Simple View and the Single Phenomenon

What would these accounts say about the two cases of interrupting that we have identified? The idea would be that the first case is intentional because I do believe that my action has the property of interrupting my partner's rehearsal, but that the second case is unintentional because I do not believe that my action has the property of interrupting my partner's rehearsal.[5]

What would the traditional account of intentional action say about the case of interrupting in the priming experiment? Well, it certainly does not look as though we could say, on the traditional account, that the experimentee interrupts the experimenter unintentionally. There is no cognitive gap (missing belief or false belief) upon which we could found the judgment that the experimentee does not intentionally interrupt. It is not as if the experimentee does not notice that the experimenter is involved in conversation with someone else.

Here it may be that the popcorn case is not equivalent. One could suggest the following: cinema-goers are intentionally eating popcorns but unintentionally eating *stale* popcorns.[6] This is modelled on Davidson's classic (1980: 84-85) example of intentionally boarding a plane headed to 'London' but unintentionally boarding a plane headed to London, Ontario. Actions can be intentional under one description but unintentional under a different one: mistakes are a paradigmatic example. But that's also the problem with this interpretation: it is not as if those cinema-goers can say that they meant to eat fresh popcorns instead of stale popcorns the way in which Davidson's character can say that she meant to board a plane headed to London, England instead of a plane headed to London, Ontario. Why not? Davidson's character can say the following: had I known that this plane was headed to London, Ontario I would not have boarded it. Can't the cinema-goers similarly say that had they known that the popcorns were stale they would not have eaten it?

View in that it is meant to go in the direction of the latter by avoiding problems with the former (the authors themselves say: "it avoids problems facing the co-called "simple view" of intentional action" (63-64), but then it explicitly talks of A-ing "with A-ing as a goal", which sounds a lot like the Simple View.

[5] I avoid discussing whether we could say the same by talking of intentions because the question of whether I intended to interrupt in the first case raises issues of *double effect* (more on this in Chapter 7); but it seems at least clear that I could not have intended to interrupt in the second case, while I at least *could* have intended to interrupt in the first.

[6] This interpretation is available whichever account of the individuation of action one endorses.

There is an important parallel between the two cases, but also an important difference. The parallel is that, in both cases, the agent could stop at any time boarding/eating. She is neither being carried forcibly onto the plane in the one scenario nor is she being force-fed in the other scenario. But while there is a bit of information that would give Davidson's character reason to stop boarding the plane (the fact that the plane is headed to London, Ontario) there is no symmetrical bit of information that should lead the cinema-goers to stop eating. The fact that the popcorn is a week old is not the relevant bit of information; it is not as if the popcorn had been secretly poisoned. That would be a parallel case, in which the cinema-goers have an external reason to stop eating (that the popcorn has been poisoned) but no internal reason to stop eating because they have no access to the relevant secret, just as Davidson's character has no access to the secret that the plane is headed to London, Ontario. Cinema-goers have access to all the information they need in that they can *taste* their popcorn. Their situation is more like the one in which someone would point out to Davidson's character that her plane is actually headed to London, Ontario for her to reply that any London will do and board the plane. London, England would have been better, but London, Ontario will do too. Fresh popcorn would have been better, but stale popcorn will do too.

Both interrupting the experimenter and eating stale popcorn do not count as unintentional actions on traditional action theories. Does that mean that these actions are intentional? The short answer is yes; the long answer is that there are action-descriptions of what the experimentees do in these experiments that would not count as intentional, but that does not mean anything because there are always alternative action-descriptions which are not intentional. So we could say that on traditional accounts 'eating one-week-old' popcorns is not intentional because the agent does not know that the popcorns are one week old; but that's the same sense in which we can say that, in going for a walk, I did not intentionally take 8743 steps, nor did I intentionally walk down 'Forest Road', just because I didn't know about these properties of my actions: that does not make my action of taking a walk unintentional, just as not knowing that the popcorns are one week old does not make my eating them unintentional.

Indeed, with the priming cases it is even more difficult to argue that they are unintentional, as there aren't any easily available unintentional descriptions: one can say, at the most, that the experimentees in the rude group did not intentionally 'interrupt more often than those in the polite group', or that the experimentees in the 'old stereotype' group did not

intentionally 'walk slower than those in the control group'; this, again, does not say anything about the intentionality of walking or interrupting.

Traditional action theories hold that agents' behaviour in these experiments is intentional; that, by the way, seems to also be the observer's intuition and the intuitions that agents themselves also have when in the debriefing they show no awareness of the role of the prime.[7] What is, then, the challenge for traditional action theory coming from these experiments? In the next section I argue that the empirical data provides a challenge to traditional action theory.

3. The experiments' challenge

We must distinguish between the claim that a combination of psychological states such as desires, beliefs, and intentions (with a specific content) is necessary for intentional action and the claim that such a combination of psychological states is sufficient for intentional action. Normally, this distinction is important with respect to the problem of deviant causal chains, where performances that are intuitively unintentional and accidental meet the conditions for intentional action. A large literature on the problem of deviance tries therefore to deal with these counterexamples in order for traditional action theories to be able to offer psychological states as sufficient conditions.[8] But this is not our concern here: our concern in distinguishing between the necessity of psychological

[7] Here I should mention that one may take evidence on so-called 'explanatory vacuum' to speak against the idea that the experimentees take themselves to be acting intentionally. Experimenters have recorded some agitation in subjects that have been primed which they have not recorded in subjects that they have consciously induced to behave in a comparable manner (Oettingen et al. 2006, Parks-Stamm et al. 2010). To stay with one of the examples I discuss here, the idea would be that those who have been primed to walk slower are afterwards agitated while those who have been explicitly warned to walk slower because, say, the floor in the corridor is wet, do not show similar levels of agitation. I won't discuss this issue in any detail here, but one could pursue this evidence in an attempt to argue that subjects do not experience themselves as acting intentionally in priming cases. But the explanatory vacuum data also has a much more important role to play for my argument: that agents are agitated speaks against the attribution of the relevant psychological states: why would people experience negative affect about what they do after having been primed if they were motivated to do it? To put it another way: if they had the relevant psychological states motivating them to act as they do, then what would explain the experienced negative affect?

[8] For a good introduction to this debate see Stout 2010. I look at this issue in detail in Chapter 4.

states and their sufficiency is that if psychological states are sufficient for intentional action, and if the examples in question are indeed intentional actions as we have argued in the previous section, then influences such as priming should either not manifest themselves in the agent's behaviour or they should manifest themselves in the agent's psychological states. This disjunctive is obviously rhetorical in so far as the first disjunct is contradicted by the experimental evidence under discussion.

This point should be distinguished by a superficially similar one, according to which, in short, traditional action theory has to explain phenomena such as priming. I think it would be unfair to ask action theory to explain why some experimentees interrupt while others don't. That's just not the job of the philosophy of action, or at least not of those in the philosophy of action who worry about reasons explanation and intentional action. I don't think that we can ask an account of what it is for an action to be intentional or of an account of what it is for something to be a reason for action to explain interruption or walking speed or improved performance in priming cases.

From the claim that we can't ask action theory to explain priming effects does not follow, though, the claim that priming effects cannot be a challenge for action theory. Because even though we can't ask an account of intentional action for an explanation of priming effects, we can ask to an account which takes the psychological states of desire, belief, or intention to be in some combination sufficient for intentional action which part of this combination of psychological states represents the priming. Namely, we can't ask an account of intentional action to explain priming effects, but we can ask an account of intentional action to rationalize those intentional actions that subjects perform after having been primed. The latter is a philosophically legitimate request to traditional accounts of intentional action, such that if these accounts cannot satisfy this request, then priming speaks against the accounts in question.

The following action-description offers a true description of what agents do after having been primed: "S interrupted rather than not interrupting". This action-description, as I have argued in the previous section, is an intentional action-description or, for short, intentional action (as intentionality, as we have also shown in the previous section, is a property of action-descriptions). What this means is that traditional accounts of intentional action can be legitimately expected to rationalize "S interrupted rather than not interrupting". If they fail to do so, then "S interrupted rather than not interrupting" will constitute a counterexample to those accounts of intentional action that fail to rationalize it.

Let us put this point in another way, which I take to be equivalent but helps illustrate my argument further: if some factor X can make a difference as to whether the agent does A or does not do A, and traditional action theory explains the agent's doing A instead of not doing A in terms of the agent's psychological states and only her psychological states (sufficiency), then factor X must be found somewhere among the agent's psychological states. And if factor X cannot be found anywhere amongst the agent's psychological states, then that shows that the agent's psychological states are not sufficient to explain why the agent does A instead of not doing A.

Here we should probably distinguish between different cases amongst the experimental evidence: what we just said, for example, seems to apply to the interrupting case but not to the walking slower case. Whether or not one interrupts seems to be different from whether one walks at 4,8km/h or at 4,5km/h. The latter seems equivalent to changing the acclivity of the floor under which agents routinely walk and record how the agent's walking speed will adapt to the new acclivity. If you are subtle enough, the change will go unnoticed.

Still, even for the walking case we can say something similar to what we said for the interrupting case: if something the agent is unaware of can change the agent's behaviour, then the agent's behaviour cannot be purely conscious behaviour. If it were purely conscious behaviour, it would not be changed by factors the agent is not aware of.[9] The difference with the interrupting case and the popcorn case is, though, that while we may demand that a reasons explanation explain us why the agent interrupted instead of not interrupting and why the agent eat popcorn instead of not eating popcorn, it may be too much to demand of a reasons explanation to tell us why the agent walked at 4,8km/h instead of 4,5km/h.

Taking the experimental evidence seriously means, I think, accepting that the only difference between those experimentees that interrupted and those that did not interrupt was the priming and its effects; similarly, it means accepting that the only difference between those who ate stale popcorns and those who didn't was their habits and their effects. Now, obviously to accept that the only differences are the priming and its effects in the one case and the habit and its effects in the other case is not the same as to accept that there is no difference in psychological states, otherwise the argument would already be settled from the start. The question which remains to be answered, then, is whether the differences

[9] At the end of this section I discuss the possibility that the psychological states which explain those behaviours are unconscious states of the agent.

constituted by priming and habits can be captured in terms of psychological states.

Let us start from the difference between those who ate stale popcorn and those who did not eat stale popcorn. An explanation in terms of psychological states would appeal to things like the belief, in one case, that the popcorn was stale against the lack of such belief in the other case. The idea would be that both agents or groups of agents had a pro attitude towards eating popcorn but one agent or group of agents also had the belief that the popcorn was stale which explains why this agent or group did not eat the stale popcorn. As the other agent or group did not have the relevant belief, they ate the stale popcorn. There are two problems with this *belief* approach: firstly, it flies in the face of what participants said afterwards about the popcorn. Namely, participants across conditions in both studies reported liking the fresh popcorn and disliking the stale popcorn (2011:5 for study 1 and 2011: 7 for study 2). So it is not as if they did not notice the difference. The cognitive relation of all participants across conditions to the popcorn did not vary with whether or not they had the habit. Secondly, if what explains why an agent or group ate the stale popcorn while the other agent or group didn't was the presence or absence of a belief about the popcorn being stale, then the eating patterns should not be sensitive to conditions such as location or method of eating – but the experiments show that they are.

The same argument would apply to trying to explain the difference between those who eat stale popcorn and those who don't in terms of their desires or other pro attitudes; if the desires of the agents would explain why they eat stale popcorn rather than not, then their eating patterns should be sensitive to a change in pro attitude rather than being sensitive, as the experiments show, to changes in environmental cues.

It could be objected to the arguments above that changes in the environmental cues can be themselves explained in terms of the agent's psychological states, in so far as a change of location is likely to result in a change of the agent's relevant psychological states. One could just say that, for example, some environments trigger (maybe even automatically) in some people a desire or an intention to eat popcorn (the cinema) while some other environments do not trigger in the same people a desire or an intention to eat popcorn (the meeting room). This is plausible: places are associated to smells, they are associated to memories, and these things may contribute to generate the relevant psychological states. But this, again, does not explain the experimental data across all conditions: it does not explain, for example, why an agent or group would eat stale popcorn when eating with her dominant hand but not eat stale popcorn when eating

BIBLIOGRAPHY

Aarts H. & Dijksterhuis A. (2003). The Silence of the Library. Journal of Personality and Social Psychology. Vol. 84, No. 1, 18–28.

Adams, F. 1986. Intention and intentional action: The Simple View. Mind&Language 1: 281–301.

Adams, F. and A. Mele (1989), 'The Role of Intention in Intentional Action', Canadian Journal of Philosophy 19, 511-31.

Alston, W. (1986), 'An Action-Plan interpretation of purposive explanations of actions', Theory and Decision 20, 275-299.

Alvarez, M. (2009), 'Actions, Thought-Experiments, and the Principle of Alternate Possibilities', Australasian Journal of Philosophy 87/1: 61-81.

Alvarez, M. (2010), Kinds of Reasons. Oxford UP.

Anscombe, G.E.M. 1957. Intention. Oxford: Basil Blackwell.

—. (1982). Medalist's Address: Action, Intention, and ´Double Effect´. Proceedings of the American Catholic Philosophical Association 56: 12-25.

Aquinas, T. (1988) Summa Theologica II-II, Q. 64, art. 7, "Of Killing". In William P. Baumgarth and Richard J. Regan, S.J. (eds.), On Law, Morality, and Politics. Indianapolis/Cambridge: Hackett Publishing Co.

Aristotle (1925), Nicomachean Ethics, translated with an introduction by David Ross, Oxford: Oxford University Press.

Armstrong D.M. (1973). Belief, truth, and knowledge (Cambridge: Cambridge UP).

Audi R. (1994). "Dispositional beliefs and dispositions to believe", Nous, 28, 419–434.

—. (1986), 'Acting for Reason', Philosophical Review 95, 511-546.

Bargh J.A., Chen M. & Burrows L. (1996). Automaticity of Social Behavior. Journal of Personality and Social Psychology Vol. 71, No. 2, 230-244.

Bargh, J.A. & Chartrand, T.L. (1999), 'The Unbearable Automaticity of Being', American Psychologist 54: 462-479.

Bargh, J.A. & Fitzsimons, G.M. (2003), 'Thinking of You: Nonconscious Pursuit of Interpersonal Goals Associated With Relationship Partners', Journal of Personality and Social Psychology 84: 148-164.

Bargh, J.A. (2006), ‘What Have We Been Priming All These Years?’, European Journal of Social Psychology, 36, 147-168.

—. (2008), ‘Free Will is Un-natural’, in Baer, J., Kaufman, J.C., & Baumeister, R.F. (eds.), Are We Free? Oxford University Press.

Bargh, J.A., Gollwitzer, P.M., Lee-Chai, A., Barndollar, K., & Trotschel, R. (2001), ‘The Automated Will: Nonconscious Activation and Pursuit of Behavioral Goals’, Journal of Personality and Social Psychology 81: 1014-1027.

Beilock S.L., Bertenthal A.M., McCoy A.M., & Carr T.H. (2004). Haste does not always make waste. Psychonomic Bulletin and Review 11: 373-79.

Beilock S.L., Carr T.H., MacMahon C., & Starkes J.L. (2002). When paying attention becomes counterproductive. Journal of Experimental Psychology: Applied 8: 6-16.

Bennett, J. (1981). Morality and Consequences. University of Utah Press.

Bermudez, J. (1995), ‘Nonconceptual Content: From Perceptual Experience to Subpersonal Computational States’, Mind and Language10: 333-369.

Bishop, J. (1989), Natural Agency. An Essay on The Causal Theory of Action. Cambridge: Cambridge University Press.

Block N. (1995). On a Confusion About the Function of Consciousness. Behavioral and Brain Sciences 18:227--47.

Brand, M. (1984), Intending and Acting. Cambridge, MA: The MIT Press.

Bratman, M. (1984), 'Two Faces of Intention', Philosophical Review 93: 375-405.

—. (1987), Intention, Plans, and Practical Reason. Cambridge, Mass.: Cambridge University Press.

Cavanaugh, T.A. (2006). Double-Effect Reasoning. Oxford UP.

Chisholm, R. (1966), ‘Freedom and Action’, in K. Lehrer (ed.), Freedom and Determinism. New York: Random House, 11-44.

Clarke, R. (2009), ‘Dispositions, Abilities to Act, and Free Will: The New Dispositionalism’, Mind 118: 323-351.

—. (2010), ‘Intentional Omissions’, Nous 44 (1), 158-177.

—. (2011), ‘Omissions, Responsibility, and Symmetry’, Philosophy and Phenomenological Research 82 (3), 594-624.

Collins, A. W. (1997), ‘The psychological reality of reasons’, Ratio, X: 108-123.

Dancy, J. (2000), Practical Reality, Oxford: Oxford University Press.

Davidson, D. (1963), 'Actions, Reasons, and Causes', Journal of Philosophy 60: 685-700.

—. (1971), 'Agency', in: R. Binkley, R. Bronaugh, and A. Marras (eds.), Agent, Action, and Reason. Toronto: University of Toronto Press, 3-37.
—. (1973), 'Freedom to Act', in Honderich, T. (ed.), Essays on Freedom and Action. Routledge and Kegan Paul, 137-56.
—. (1978), 'Intending', in Y. Yovel (ed.), Philosophy of History and Action, Jerusalem: The Magnes Press, The Hebrew University, 83-102.
—. (1980), Essays on Actions and Events, Oxford: Oxford University Press.
—. (2005). Truth, Language, and History. Oxford UP.
Davis, N. (1984). The Doctrine of Double Effect: Problems of Interpretation. Pacific Philosophical Quarterly 65: 107-123.
Dennett, D.C. (1991), Consciousness Explained. London: Penguin.
Di Nucci, E. & McHugh, C. (2006) (eds.), Content, Consciousness, and Perception. CSP.
Di Nucci, E. (2008), Mind Out of Action. VDM Verlag.
—. (2009a), Simply, false', Analysis 69 (1), 69-78.
—. (2009b), Abortion: Strong's counterexamples fail. J Med Ethics 35:304-305.
—. (2009c), On how to interpret the role of the future within the abortion debate. J Med Ethics 35:651-652.
—. (2010a). Rational constraints and the Simple View. Analysis 70: 481–86.
—. (2010b), Refuting a Frankfurtian Objection to Frankfurt-Type Counterexamples, Ethical Theory and Moral Practice 13 (2), 207-213.
—. (2011a), Frankfurt counterexample defended, Analysis 71 (1), 102-104
—. (2011b), Frankfurt versus Frankfurt: a new anti-causalist dawn, Philosophical Explorations 14 (1): 1-14.
—. (2011c), Sexual Rights and Disability, Journal of Medical Ethics 37 (3): 158-161.
—. (2011d), Automatic Actions: Challenging Causalism, Rationality Markets and Morals 2 (1): 179-200.
—. (2012a), Knowing Future Contingents, Logos & Episteme 3 (1): 43-50.
—. (2012b), Double Effect and Assisted Dying. British Medical Journal (letter, 7.2.2012).
—. (2012c), Priming Effects and Free Will. International Journal of Philosophical Studies 20(5): 725-734.
—. (forthcoming a), Fathers and Abortion. The Journal of Medicine & Philosophy.

—. (forthcoming b), Killing Foetuses and Killing Newborns. Journal of Medical Ethics.
—. (forthcoming c), Self-Sacrifice and the Trolley Problem. Philosophical Psychology.
—. (forthcoming d), Withdrawing artificial nutrition and patients' interests. Journal of Medical Ethics.
—. (forthcoming e), Double Effect and Terror Bombing. GAP.8 Proceedings.
—. (forthcoming f), Embryo Loss and Double Effect. Journal of Medical Ethics.
—. (forthcoming g), Habits, Nudges, and Consent. American Journal of Bioethics.
—. Ethics without Intention (book manuscript).
Dreyfus, H. & Dreyfus, S. (1984), 'Skilled Behavior: The Limits of Intentional Analysis', in Lester, E. (ed.), Phenomenological Essays in Memory of Aron Gurwitsch. The University Press of America.
Dreyfus, H. (1988), 'The Socratic and Platonic Bases of Cognitivism', AI & Society 2: 99-112.
—. (2005), 'Overcoming the Myth of the Mental', Topoi 25 (1-2), 43-49.
Eilan, N. (eds.), Agency and Self-Awareness. Oxford: Clarendon Press.
Fara, M. (2008), 'Masked Abilities and Compatibilism'. Mind, 117, pp. 843–65.
Fischer, J.M. & Ravizza, S.J. (1998), Responsibility and Control. Cambridge UP.
Fischer, J.M. (1982), 'Responsibility and Control', The Journal of Philosophy 79/1: 24-40.
—. (1994), The Metaphysics of Free Will. Blackwell.
—. (1999), 'Recent Work on Moral Responsibility', Ethics 110/1: 93-139.
—. (2008), 'Freedom, Foreknowledge, and Frankfurt: a reply to Vihvelin', Canadian Journal of Philosophy 38: 327-342.
FitzPatrick, W.J. 2009. Thomson's turnabout on the trolley. Analysis 69 (4): 636-43.
Foot, P. 1967. The problem of abortion and the doctrine of the double effect. Oxford Review 5: 5–15.
Frankfurt, H. (1969), 'Alternate Possibilities and Moral Responsibility', Journal of Philosophy 66: 829-39.
—. (1978), 'The Problem of Action', American Philosophical Quarterly 15, 157-162.
Garcia, J.L.A. 1990. The intentional and the intended. Erkenntnis 33: 191–209. Ginet, C. 1990. On Action. Cambridge: Cambridge University Press.

Gendler, Tamar Szabó (2008a), “Alief and belief”, Journal of Philosophy, 105: 634–663.
—. (2008b), “Alief in action, and reaction”, Mind and Language, 23: 552–585.
Gert, B. (1998), Morality: its nature and justification, Oxford: Oxford University Press.
Gertler, Brie (forthcoming). ‘Self-Knowledge and the Transparency of Belief’, in A. Hatzimoysis (ed.) Self-Knowledge. Oxford: Oxford.
Gigerenzer G. (2007). Gut Feelings. Penguin.
Ginet, C. (1990), On Action. Cambridge: Cambridge University Press.
Goldie, P. (2000), ‘Explaining expressions of emotions’, Mind 109: 25-38.
Gorr, M. and T. Horgan. 1982. Intentional and unintentional actions. Philosophical Studies 41: 251–62.
Grice, H. P. (1971), ‘Intention and Uncertainty’, Proceedings of the British Academy 57, 263-79.
Hampshire, S. (1959), Thought and Action. London: Chatto and Windus.
Hart, H.L.A. (1967). Intention and Punishment. Oxford Review 4.
Hauser, M. 2006. Moral Minds. New York: HarperCollins.
Hornsby, J. (1980), Actions, London: Routledge & Kegan Paul.
Huebner, B. & Hauser, M. 2011. Moral judgments about altruistic self-sacrifice: When philosophical and folk intuitions clash. Philosophical Psychology 24 (1): 73-94.
Hume, D. (1978), Treatise of Human Nature, (Nidditch edition), Oxford: Oxford University Press.
Hunter, David (2009). ‘Belief, Alienation, and Intention’, unpublished manuscript.
Hursthouse, R. (1991), 'Arational Actions', Journal of Philosophy 88 (2), 57-68.
Iyengar S.S. & Lepper M.R. (2000). When choice is demotivating. Journal of Personality and Social Psychology 79: 995-1006.
James, W. (1890), The Principles of Psychology, London: Macmillan.
Kamm, F. (1996). Morality, Mortality (vol. II). Oxford UP.
—. (2007). Intricate Ethics: Rights, Responsibilities, and Permissible Harm. New York: Oxford University Press.
Kahneman, D. (2011), Thinking, Fast and Slow. Penguin.
Kant, I. (1785). The Groundwork of the Metaphysics of Morals.
Klein, G.A. (1999). Sources of Power. MIT Press.
Liao, S.M., Wiegmann, A., Alexander, J., Vong, G. Putting the Trolley in Order: Experimental Philosophy and the Loop Case. Philosophical Psychology (forthcoming).

Libet, B. (1985), ‘Unconscious cerebral initiative and the role of conscious will in voluntary action’, Behavioural and Brain Science 8: 529-566.
Lowe, J. (1999), ‘Self, Agency, and Mental Causation’, Journal of Consciousness Studies 6, 225-39.
Lycan W.G. (1986) “Tacit belief”, in R.J. Bogdan, ed., Belief: Form, content, and function (Oxford: Clarendon), 61–82.
Macrae, C.N. & Johnston, L. (1998), ‘Help, I Need Somebody: Automatic Action and Inaction’, Social Cognition 16: 400-417.
Mandelbaum E. (2012). Against Belief. Philosophical Studies (forthcoming).
Mangan, J.T. (1949). An Historical Analysis of the Principle of Double Effect. Theological Studies 10: 41-61.
McCann, H. 1991. Settled objectives and rational constraints. American Philosophical Quarterly 28: 25–36.
—. 2010. Di Nucci on the Simple View. Analysis 70: 53–59.
—. 2011. The Simple View again: a brief rejoinder. Analysis 71(2): 293-95.
McDowell, J. (1978), 'Are Moral Requirements Hypothetical Imperatives?', Proceedings of the Aristotelian Society, Supplementary Volume 52, 13-29.
—. (1994), Mind and World, (with a new introduction by the author: 1996), Cambridge, MA: Harvard University Press.
McKenna, M. (1997), 'Alternative Possibilities and the Failure of the Counter-Example Strategy', Journal of Social Philosophy 28 (3): 71-85.
Melden, I. (1961), Free Action, London: Routledge Kegan & Paul.
Mele, A. (1992), Springs of Action, Oxford: Oxford University Press.
—. (1997), ‘Passive Action’, in: G. Holmstrom-Hintikka and R. Tuomela (eds.), Contemporary Action Theory. Dordrecht: Kluver, 135-143.
—. (1997), Philosophy of Action. Oxford UP.
Mele, A. and Moser, P. K. (1994), 'Intentional Action', Nous 28: 39-68.
Moya, C. (1990), The Philosophy of Action, Cambridge: Polity Press.
Nagel, T. (1986). The View from Nowhere. Oxford University Press.
Neal D.T., Wood W., Wu M., Kurlander D. (2011). The Pull of the Past. Personality and Social Psychology Bulletin XX(X): 1-10.
Newman G.E., Diesendruck G., Bloom P. (2011). Celebrity Contagion and the Value of Objects. The Journal of Consumer Research (forthcoming).
Norman, D.A. & Shallice, T. (1986), ‘Attention to Action: willed and automatic

O'Shaughnessy, B. (1980), The Will: Volume 2, A Dual Aspect Theory, Cambridge: Cambridge University Press.
Oettingen G., Grant H., Smith P.K., Skinner M., Gollwitzer P.M. (2006). Nonconscious goal pursuit: Acting in an explanatory vacuum. Journal of Experimental Social Psychology 42: 668-75.
Otsuka, M. (1997). Kamm on the Morality of Killing. Ethics 108 (1): 197-207.
—. (2008). Double-Effect, Triple-Effect and the Trolley Problem. Utilitas 20: 92-110
Parks-Stamm E.J., Oettingen G., Gollwitzer P.M. (2010). Making sense of one's actions in an explanatory vacuum. Journal of Experimental Social Psychology 46: 531-42.
Perner, J. (2003), 'Dual control and the causal theory of action', in Roessler, J. &
Petrinovich, L., and O'Neill, P., (1996). Influence of wording and framing effects on moral intuitions. Ethology and Sociobiology 17: 145-171.
Pippin, R. B. (2008), Hegel's Practical Philosophy, Cambridge: Cambridge University Press.
Pollard, B. (2003), 'Can Virtuous Actions Be Both Habitual and Rational?', Ethical Theory and Moral Practice 6, 411-425.
—. (2005), 'Naturalizing the Space of Reasons', in International Journal of Philosophical Studies 13(1), 69-82.
—. (2006), 'Actions, Habits, and Constitution', Ratio 19, 229-248.
—. (2006), 'Explaining Actions with Habits', American Philosophical Quarterly 43: 57-68
Price, H. H. (1969). Belief. London: George Allen & Unwin.
Ramachandran, V.S. and W. Hirstein. 1998. The perception of phantom limbs. Brain 212: 1603–30.
Rietveld, E. (2008), 'Situated Normativity: The Normative Aspect of Embodied Cognition in Unreflective Action', Mind 117, 973-1001.
Rowbottom, Darrell P. (2007). ' "In-Between Believing" and Degrees of Belief', Teorema 26,pp. 131–7.
Ryle, G. (1949), The Concept of Mind. London: Penguin.
Sartorio, C. (2005), 'A new asymmetry between actions and omissions', Nous 39: 460-482.
—. (2009), 'Omissions and Causalism', Nous 43: 513-530.
Schlosser, M.E. (2007), 'Basic deviance reconsidered', Analysis 67 (3): 186-194.
—. (2010), 'Bending it like Beckham: movement, control, and deviant causal chains', Analysis 70 (2): 299-303.

Schwitzgebel E. 2010a "Acting Contrary to Our Professed Beliefs, or the Gulf Between Occurrent Judgment and Dispositional Belief" Pacific Philosophical Quarterly, 91, 531-553.
Schwitzgebel E. 2010b Belief. Stanford Encyclopedia of Philosophy.
Searle J.R. (1992) The rediscovery of the mind (Cambridge, MA: MIT).
Searle, J. (1983), Intentionality. Cambridge: Cambridge University Press.
Sinnott-Armstrong, W. (2008). Framing Moral Intuitions in W. Sinnott – Armstrong (Ed.) Moral Psychology, Volume 2: The Cognitive Science of Morality, (pp. 47-76). Cambridge, MA: MIT Press.
Smith, M. (1987), 'The Humean Theory of Motivation', Mind 96, 36-61.
—. (1996), The Moral Problem. Cambridge MA: Harvard University Press.
—. (2003), 'Rational Capacities, or: How to Distinguish Recklessness, Weakness, and Compulsion', in Stroud, S. & Tappolet, C. (eds.) Weakness of Will and Practical Irrationality. Oxford UP.
Sommers, Fred (2009). 'Dissonant Beliefs', Analysis 69, pp. 267–74.
Stout, R. (1996), Things that happen because they should. Oxford: Oxford University Press.
—. 2005. Action. McGill: Queen's University Press.
—. (2010), 'Deviant Causal Chains', in O'Connor, T. & Sandis, C. (eds.), A Companion to the Philosophy of Action. Blackwell: 159-165.
Stoutland, F. (1985), 'Davidson on Intentional Behaviour', in: E. LePore and B.P. McLaughlin (eds.), Actions and Events, Oxford: Basil Blackwell, 44-59.
Strawson, P.F. (1962), 'Freedom and Resentment', Proceedings of the British Academy 48: 1-25.
Sverdlik, S. 1996. Consistency among intentions and the 'Simple View'. Canadian Journal of Philosophy 26: 515–22.
Swain, S., Alexander, J. and Weinberg, J. (2008). The Instability of Philosophical Intuitions: Running Hot and Cold on Truetemp. Philosophy and Phenomenological Research 76: 138-155.
Tanney, J. (1995), 'Why Reasons May Not be Causes', Mind & Language 10, 103-126.
Tannsjo, T. (2009), 'On deviant causal chains – no need for a general criterion', Analysis 69: 469-473.
Thalberg, I. (1977), Perception, Emotion, and Action, Oxford: Basil Blackwell.
—. (1984), 'Do our intentions cause our intentional actions?', American Philosophical Quarterly 21, 249-260.
Thaler R.H. & Sunstein C.R. (2008). Nudge. Caravan.

Thomson, J.J. (1977), Acts and Other Events, Ithaca, N.Y.: Cornell University Press.
—. 1976. Killing, letting die, and the trolley problem. The Monist 59: 204–17.
—. 1985. The trolley problem. The Yale Law Journal 94: 1395–415.
—. 2008. Turning the trolley. Philosophy and Public Affairs 36: 359–74.
Van Inwagen, P. (1983), An Essay on Free Will. Oxford UP.
Vihvelin, K. (2000), 'Freedom, Foreknowledge, and the Principle of Alternate Possibilities', Canadian Journal of Philosophy 30: 1-24.
—. 2004: 'Free Will Demystified: A Dispositional Account'. Philosophical Topics, 32, pp. 427–50.
—. (2008), 'Foreknowledge, Frankfurt, and Ability to do otherwise: reply to Fischer', Canadian Journal of Philosophy 38: 343-372.
von Wright, G.H. (1971), Explanation and Understanding. Cornell UP.
Wedgwood, R. (2011). Defending Double Effect. Ratio (forthcoming).
Wiegmann, A., Okan, Y., Nagel, J., & Mangold, S. (2010). Order Effects in Moral judgment. In S. Ohlsson & R. Catrambone (Eds.), Proceedings of the Thirty-Second Annual Conference of the Cognitive Science Society (pp. 2111-2116). Austin, TX: Cognitive Science Society.
Wilson, Timothy D., Samuel Lindsey, and Tonya T. Schooler (2000), "A model of dual attitudes", Psychological Review, 107: 101–126.
Withehead, A.N. (1911), An Introduction to Mathematics, New York: Holt.
Wittgenstein, L. (1969), On Certainty. Basil Blackwell.
Wood W. (2012). On Ruts and Getting Out of Them. Science 336 (6084): 980-81.
Woodward, P.A. (ed.) (2001). The Doctrine of Double Effect. University of Notre Dame Press.
Zhu, J. (2004), 'Passive Action and Causalism', Philosophical Studies 119: 295-314.
Zimmerman A. (2007), "The nature of belief", Journal of Consciousness Studies, 14(11): 61–82.