

Responsibility and Health

Responsibility and Health

Explanation-Based Perspectives

Karl Persson de Fine Licht



UNIVERSITY OF GOTHENBURG
PHILOSOPHY, LINGUISTICS & THEORY OF SCIENCE

Doctoral dissertation in Practical Philosophy
University of Gothenburg, 2014

© Karl Persson de Fine Licht, 2014

Typesetting: Adobe Garamond

Cover design: Thomas Ekholm and Karl Persson de Fine Licht

ISBN 978-91-628-8887-9 (Hard Copy)

ISBN 978-91-628-8888-6 (Electronic Copy)

Available at: <http://hdl.handle.net/2077/34421>

Distribution: The Department for Philosophy, Linguistics, and Theory of Science, Box 200,
405 30 Gothenburg

To my loving family...

Abstract

There is a growing trend of holding people responsible for their lifestyle-based diseases. For example, policymakers as well as researchers have argued that medical conditions caused by smoking, overweight, or extreme sports should be given lower priority in publicly funded healthcare. This view on resource allocation is grounded in a common belief that people are responsible for these diseases and that it is fair to hold them responsible because of this.

It is far from clear, however, that we are responsible for anything we do, never mind our ill health. A wide range of skeptical arguments seem to show that the notion of responsibility is misplaced. Though these arguments are controversial, they have by no means been successfully refuted. Furthermore, even if these skeptical arguments were shown to be invalid, there are additional, more mundane, excusing conditions that might relieve people of responsibility, such as those from nonculpable ignorance and addiction.

Responsibility and Health: Explanation-Based Perspectives presents four papers that address this problem from a new angle. The first two papers show that the discussion about the conditions for responsibility should be reframed, and they provide new grounds for rejecting the skeptical arguments. The third paper discusses the validity of the skeptical arguments further and shows that even though healthcare personnel should reject the skeptical arguments, policymakers should not. The fourth paper examines whether the obese are responsible for being obese and finds that they are not. This, in conjunction with other arguments, gives us reason to prioritize overweight surgery.

Taken together, the four papers offer new insights in the discussion about responsibility and health and a promising avenue for making progress in the responsibility debate.

List of Papers

- Paper 1** Björnsson, G.; Persson, K. (2012) The Explanatory Component of Moral Responsibility. *Noûs* 46 (2) pp. 326-354.
- Paper 2** Björnsson, G.; Persson, K. (2013) A Unified Empirical Account of Responsibility Judgments. *Philosophy and Phenomenological Research* 87 (3): 611-639.
- Paper 3** Persson, K. (2012) The Right Perspective on Responsibility for Ill Health. *Medicine, Health Care, and Philosophy* 16 (3): 429—421.
- Paper 4** Persson, K. (2012) Why Bariatric Surgery Should be Given High Priority: An Argument from Law and Morality. *Health care analysis*. DOI: 10.1007/s10728-012-0216-1.

Acknowledgments

It's Boxing Day. Manchester United has just won over Hull after recovering from conceding two early goals, and I am beginning to see the end of my work on this thesis. These four years have been frustrating, disheartening, and full of gloom. But they have also been inspiring, fun, and extremely developing which, in the end, have made it all worthwhile.

Being in one place, doing one thing, for such a long period of time, there are, of course, many individuals (and institutions!) to thank. To begin with, I would like to thank my main advisor, Gunnar Björnsson, for his constant encouragement and his thorough comments on most of my works. Without his input, my dissertation would be a bleak version of what it is now. It is even a strong possibility that it would not be a dissertation at all. I would also like to thank him for bringing me on as a co-author very early in my career, showing me the ropes on how it's done in the big leagues. (He is also a master of acronyms, a hidden talent I've found to be very useful).

I would also like to thank my (co-)advisor, Christian Munthe, for commenting on much of my work, but also for bringing me into interesting projects and dealing with the practical matters regarding my graduate school. Ingmar Persson, my first main advisor, should also be mentioned for being inspiring through the amount and quality of his own work, but also for comments on my early contributions. Perhaps I also should mention Linus Broström in this context; he did an excellent job as an opponent on my final seminar.

I would also like to thank my current and precious roomies, Joakim Sandberg, Ragnar Francén Olinder, and Sandra Lindgren for their witticisms and sometimes more serious talks, but also for reading parts of my work. Especially Joakim I would like to thank in this latter regard. I would also like to thank John Eriksson for a lot of invigorating discussions about sports and for dragging me to the action park where I fractured both of my elbows while skateboarding. Gustaf Nelhans is also someone I would like to send a big thanks to, this for helping me with the practical stuff when finishing my dissertation.

Looking outside the philosophy department, I would like to thank my wife, Jenny de Fine Licht. Many times, when people are giving thanks to their loved ones, it is for their support and for “being there.” This is, of course, something I am grateful for as well, but in addition, she has actually read many of the things I have done, papers, research proposals, and so on, and given me a lot of help making these things much better. To have such a wife, I am very grateful. Another person I would like to thank is Göran Duus-Otterström, for a great many things. When I first met Göran, I studied political science and he couldn’t read my handwriting, so I had to read my paper out loud. It didn’t start well, I couldn’t understand what I had written, but it got better. Since then he has read much of the stuff I have written, adding good comments, and he has been an inspiring lunch partner. I would also like to thank R. Jay Wallace for inviting me to Berkeley, where I wrote two of the papers in this dissertation.

Of course, there are many more that have helped me in different ways, making this journey much easier. I will here list some of them in no particular order: Ida Hallgren, Thomas Hartvigsson, Alva Stråge, Caj Strandberg, Susanna Radovic, Erik Joelsson, Lena Eriksson, Daniela Cutas, Max Fonseca, Jakob Forsbäck, Sari Nauman, Bryan Cockrell, and Julian Theis.

There is also a wide range of foundations and departments that have contributed to my conference visits, my stay abroad, etc. These are the following: Anders Karitz Foundation, Bo Samuelssons PhD-student Exchange Foundation to Berkeley, The (old) Philosophy Department, FLoV, my graduate school of Environment and Health, The Nordic Committee for Bioethics, Paul & Marie Berghaus Donation Fund, and SUN, the Central European University’s stipend foundation. I would also like to thank Springer and Wiley for letting me reprint the papers published in their journals and the multi-disciplinary graduate school, Environment and Health, for in part funding my studies.

Last, but not least, I would like to thank my mother (tack mor!) and sister for their support and encouragement and my son for distracting me and letting me be home from work for nine months.

Karl Persson de Fine Licht, Gothenburg, Sweden, December 2013.

Content

1. Introduction	13
2. Questions to be answered and strategies that can be used	16
3. The debate thus far and summary of the four papers	20
3.1 Responsibility, heteronomy, and luck	20
3.1.1 Views on responsibility	20
3.1.2 The argument(s) from heteronomy	23
3.1.3 The argument(s) from contrastive luck	29
3.1.4 Summary and outlook	36
Paper 1: The Explanatory Component of Moral Responsibility	37
3.2 Responsibility and experimental philosophy	41
3.2.1 What is experimental philosophy?	41
3.2.2 Folk intuitions and responsibility judgments	44
3.2.3 Explanations of diverging results and why they fail	54
3.2.4 Summary and outlook	57
Paper 2: A Unified Empirical Account of Responsibility Judgments	59
3.3 Responsibility, accountability, and ill health	62
3.3.1 Some introductory clarifications	62
3.3.2 Responsibilitarianism	68
3.3.3 Utilitarianism	78
3.3.4 Egalitarianism	84
3.3.5 Autonomy views	88
3.3.6 Greatest need, beneficence, non-maleficence	96
3.3.7 Summary and outlook	97
Paper 3: The Right Perspective on Responsibility for Ill Health	98

Paper 4: Why Bariatric Surgery Should be Given High Priority: An Argument From Law and Morality	101
4. Results, implications, and further research	104
4.1 Results	104
4.2 Implications	113
4.2.1 Theoretical implications	113
4.2.2 Policy implications	121
4.3 Further research	129
References	132
Papers I–IV	142
Svensk sammanfattning	147

1. Introduction

There is a growing trend of holding people responsible for their own ill health. This is seen in most parts of the western world from Europe to the United States (Persson 2012) and occurs in a number of ways and on a number of grounds. For example, people are held financially liable when being reckless with their health, because it is commonly believed that this will give them incentive to do otherwise and hence create a population with better health. Similarly, people are given a lower priority in priority-setting contexts because it seems unfair to give those who are responsible for their own ill health the same priority as those who are completely “innocent.” There are numerous other types of examples.

It is easy to see where the trend originates, at least when it comes to questions about liability and priority setting. Simply put, we can do more for people each day, but this increases the costs. This is often illustrated by the rise in the cost of medications against cancer and rheumatoid arthritis. Take one example, a new cancer drug, cetuximab, which is used to treat non-small cell lung cancer. According to a recent study, 18 weeks of treatment give the patient 1.2 extra months of life (Fojo and Grady 2009). An average cost for this period is \$80,000 and for a year is \$800,000. If every American in need (all 550,000 of them) were to receive this drug, it would cost approximately \$440 billion annually, which is about the same as Sweden’s GDP for 2012 (\$500 billion). Another example is the new biological drugs against rheumatoid arthritis. According to the British National Institute for Health and Excellence (NICE), the cost for these drugs is about £9500 a year in comparison with £450 a year for conventional treatment.¹ That is, it is an increase in cost by about 21 times.

The reason why these new drugs cost so much money in comparison with the old ones is that it costs much more to develop the new drugs than it did before. There is no reason to believe that the development costs will sink in

¹For numbers and calculations see:
http://www.nice.org.uk/media/BB8/C9/biological_therapies_PDF.pdf

the future. Instead, it is probable that they will continue rising. This increase in cost will make priority setting even more difficult and will mean that someone or some group has to be given lower priority. Here it is easy to point the finger at those who have been negligent with their health. They should pay and/or be given low priority, and not those who have been careful trying to retain good health.

Is the trend of holding people responsible for their own ill health justifiable on normative grounds? To answer this question, many people would say, we must answer at least two questions: are people substantially responsible for their own ill health, and should we hold them responsible for it?

First, it is not clear whether people are substantially responsible for anything, never mind their ill health. There are many forms of responsibility such as causal, legal, attributable, etc.² To be substantially responsible for something is to be responsible in a strong, desert implying sense. That is, when people are substantially responsible for something, we thereby have a *pro tanto* reason to hold them responsible for it, simply on the ground that they are responsible. However, according to many, we cannot be responsible in this strong sense, and because of this, it would be unfair to hold people accountable for their own ill health. It is just bad luck that they came to be unhealthy rather than healthy. Second, it is not clear that we should hold people responsible when they are substantially responsible (responsible, for short) for their own ill health. It might be thought that other values trump those of giving people what they deserve, such as helping those with the greatest needs or doing what has the most beneficial effects. The converse position is also possible: one might think that even though people are excused, we still have reason to hold them accountable for their ill health, because this has beneficial effects, for example.

Thus the overarching question in this dissertation is whether people are substantially responsible for their ill health and whether we should hold them responsible for it. The rest of this introductory summary chapter will specify the questions to be asked, give background to the questions, answer the ques-

² There are many different ways in which we use the term “responsible.” See for example Hart 1968: 211, Watson 1996, Schmidt 2009, Fischer and Tognazzini 2011.

tions by summarizing my results, and say something about what more there is to be done in the area of responsibility, accountability, and health.

2. Questions to be answered and strategies that can be used

In the previous section, I distinguished the following questions:

- (i) Under what conditions are we substantially responsible?
- (ii) Should we hold people accountable for their own ill health?

When we try to answer questions like (i) and (ii), we may either start “top-down” or “bottom-up” (see for example Wolf 2011). Starting top-down, you discern relevant existing theories of responsibility and normative considerations and then try to show which one(s) is correct. After that, you may try to make the theory and/or consideration(s) sufficiently determinate in order to be able to say something substantial about concrete cases. For example, when you try to analyze whether we should hold the obese responsible for their ill health, you would first examine what relevant normative theories there are, such as luck-egalitarianism, utilitarianism, or sufficientarianism. Then you argue that one of these is correct, and you try to develop that theory so you can say something about the plight of the obese.

Furthermore, if your theory of responsibility and your normative consideration(s) are sensitive to facts about how people actually are, how the world works, and so on, then you must also take into account the empirical work that has been done and perhaps pursue further empirical studies relevant to your theories. What these empirical facts might be can also differ, depending on the theory. For example, total-view utilitarians might be interested in radically different things than luck-egalitarians. Utilitarians are interested in maximizing the total sum of well-being and are thereby most often interested in what effects different measures will have in the future. Luck-egalitarians, on the other hand, are interested in distributing resources after responsibility desert. This means that they are interested in how the state of affairs has come about and are thereby interested in the past. This also means that they would

require radically different empirical studies to be able to say anything interesting about the questions posed in this dissertation.

A less ambitious and more commonly used top-down approach is to show that all relevant theories lead to the same conclusion and to use cases where the theories and considerations do not have to be that determinate in order to yield interesting results. In some cases, the empirical evidence is clear enough to lead to conclusions, even though the normative considerations or the responsibility theories are not that determinate. Take, for instance, the case of in vitro fertilization (IVF). In such cases, some people argue, the creation of new, happy lives and all that comes with them is so valuable that all the relevant normative considerations imply that we should give IVF a high priority in priority-setting contexts (Tännsjö 2006).

Another way of trying to answer questions such as one of those posed in this dissertation, whether or not we should hold people responsible for their own ill health, is to go “bottom-up.” We begin within the actual discussion of the subject, and as philosophers, we can, for instance, try to make the discussion clearer, see if the arguments used are coherent, look into whether the normative arguments are robust, and so on. Hopefully, by this means, one can reach some sort of conclusion after excluding poor arguments and making the discussion clearer.

The difference, then, between the “top-down” and the “bottom-up” is that in the “bottom-up” method, you do not bring the conceived normative theories with you into the discussion and there apply them to see what results they might yield. Instead, you try to disentangle and dismiss some of what has actually been said in the discussion. What I will do in this dissertation is what most philosophers do, namely, use both the top-down and bottom-up strategies. I will not follow the most ambitious top-down route, where one shows that one’s favored theory is the only one we have reason to believe in. Instead, I will try to make progress with regard to questions (i) and (ii) through answering four questions by means of the bottom-up approach, discussing the questions being posed in the different debates and using the less ambitious top-down approach that takes the conceived normative considerations into the discussion to see whether they yield the same results. First, why do many of us judge people not to be responsible when confronted with the classical

skeptical arguments from heteronomy and luck? In answering this question, I hope to move the discussion forward, relieving some of the stalemates by putting the discussion in another light. However, this will not lead to a theory on responsibility, and it will not be the final say in the debate concerning the conditions for responsibility. But when we have a better understanding of why some of the most famous skeptical arguments make us tick, we will have come a much longer way toward solving these questions. We will be able to determine the evidential value of these intuitions, whether they are caused by different biases, and so on.

Second, can we find a unified theory explaining the divergent and perplexing responsibility judgments of people in general? To say something more informed about why we react the way we do when confronting different cases and arguments, we need to know in what way non-philosophers (“the folk”) react and what other hypotheses there are that explain these judgments. In recent years, experimental philosophers have conducted numerous studies on folk intuitions of responsibility, and the results have often been perplexing. I will try to give a unified answer to why people react in these seemingly disparate ways. When we have such an answer, and that answer is the same as the answer to why philosophers react the way they do when confronting skeptical arguments, we will have come even further in resolving different forms of stalemates in the debate on conditions for responsibility. This is because such a hypothesis would be very strongly supported.

The third question I will answer concerns what perspective healthcare personnel (broadly construed) and policymakers should take on people’s responsibility for their own ill health. I argue that the conditions for responsibility should be defended on normative or metaethical grounds. Taking the normative route, we ask whether we *should* perceive the skeptical arguments as valid, that is, whether we should perceive people as links in chains (i.e. dependent variables). What I seek to do is find out whether healthcare personnel and policymakers should have a certain disposition to, or a rule of thumb for, perceiving people as responsible for their own ill health. Of course, this approach opens a range of exceptions. For example, it might be argued that people in some profession among healthcare personnel should perceive people as responsible even though, we might assume, healthcare personnel should in gen-

eral perceive their patients as excused. Still, having some sort of overarching idea about what to do is very helpful.

The fourth and last question I will discuss is much more concrete than the others and is about whether bariatric surgery should receive a high priority. This question is interesting in its own right, because bariatric surgery is an expensive procedure, producing what seem to be good results, and is being used more frequently but which could be used much more often. Strangely enough, no one has made a substantial argument for either side in this issue; I attempt to fill that significant lacuna. Furthermore, an answer to whether we should give bariatric surgery high priority or whether we should hold the obese accountable by giving them low priority says something about whether the obese are responsible for being obese, whether they have great needs, what alternatives there are for them to lose weight, and so on. Accordingly, answering the narrow question about bariatric surgery could also lead to some progress in other areas relevant to the more general question of whether we should hold people responsible for their own ill health.

To sum up: the overarching questions are these:

- (i) Under which conditions are we substantially responsible?
- (ii) Should we hold people accountable for their own ill health?

The somewhat more precise questions springing out of (i) and (ii) are these:

- 1) Why do we experience skeptical intuitions when being confronted with cases of heteronomy and moral luck, and do these intuitions have any evidential value?
- 2) How can we explain the perplexities found in experimental philosophy, and can we give them a unified explanation?
- 3) Should healthcare personnel (broadly construed) and policymakers in general perceive the skeptical arguments to be valid?
- 4) Should we give bariatric surgery high priority?

Before I give my own answers to these questions in the form of summaries of my papers, I will provide a theoretical overview of what has been said in the debate thus far.

3. The debate thus far and summary of the four papers

In this section, I will give a theoretical background to my four papers and provide an overview of previous research that provides the context for the answers I provide. I will say something about the debate about the skeptical arguments—the heteronomy and luck arguments—and then I will say something about the discussion of whether we should hold people responsible for their own ill health. After each section (i.e., 3.1, 3.2, 3.3), I will summarize the paper/papers relevant to each section.

3.1 RESPONSIBILITY, HETERONOMY, AND LUCK

To say something more substantial about the discussion of responsibility, it might be a good idea to begin by taking a closer look at the standard views on responsibility. This is what I am going to do first. After that I will relate these different theories to the two groups of skeptical arguments of heteronomy and contrastive luck. Last, I will give a summary of my paper where I explain what my contribution is and how it relates to the debate at hand.

3.1.1 Views on responsibility

Broadly speaking, there are four common positions on moral responsibility: hard incompatibilist (Strawson 1986, Smilansky 2000, Pereboom 2001), soft incompatibilist (Kane 1996, O'Connor 2000, Clarke 2003),³ compatibilist (Wallace 1994, Dennett 2003, Nelkin 2011), and what we might call “indeterminates” (Mele 1996, 2006).⁴ Roughly speaking, compatibilists believe that responsibility is compatible with our actions being determined by factors outside our control, while incompatibilists do not. Soft incompatibilists and

³ Soft incompatibilists are often called “libertarians.” Since there is a risk of confusing these with those who are libertarian with regard to distributive justice, I will continue calling them “soft incompatibilists.”

⁴ With regard to free will, there are also hard compatibilists (Hobart 1934, Persson 2005, Parfit 2012). They believe, among other things, that even though we have free will, we are not substantially responsible.

compatibilists believe that there are no fundamental problems with people being responsible, while hard incompatibilists believe the opposite, that these problems are insurmountable.

At least three of these positions, soft and hard incompatibilist and compatibilist, can be further divided into two categories: leeway and source. Those from the leeway category believe that to be responsible for what we do, we have to have alternative possibilities, that is, that we have to be able to do or act otherwise. This view can be held both by compatibilists (Lewis 1981, Berofsky 1995, Smith 2003) and incompatibilists (Inwagen 1983, Ekstrom 2000). Those from the source category believe that the agent has to be the source of her actions in order to be responsible for the actions. These views can also be both compatibilist (Dennett 1984, Nelkin 2011) and incompatibilist (Strawson 1986, Kane 1996, Pereboom 2001). I will now discuss the leeway and source views in some detail and show how they relate to the skeptical arguments.

Both leeway compatibilists and incompatibilists believe that we have to be able to act otherwise in order to be substantially responsible. Certainly, they do not agree about how this condition should be spelled out in further detail, but they do agree about the necessity of this sort of condition, and it is easy to see why. Take Ann, who is suffering from being a kleptomaniac. Assume that she suffers from an extreme compulsion disorder so that she cannot help herself from shoplifting. When she is inside a store, she always tries very hard not to steal, but she always fails. Since Ann is a good person, she always goes back with the items she has stolen, and now she always warns the store personnel about her condition beforehand so that they can see to it that she gets help not to steal. Sometimes they pick up her groceries while she is waiting at the cash register, and at other times, they are sure to check her belongings when she is leaving. Presumably, few people think Ann is responsible, or at least not fully responsible, for her acts of stealing. Why is this so? One plausible explanation is that she could not avoid stealing. She could not act otherwise.⁵

⁵ Our residual inclination to still judge her to be somewhat responsible might be traced to the fact that we have a hard time picturing someone who has no ability to do otherwise. The idea that we are all able to do otherwise is simply so deeply ingrained in our conception of ourselves that these conditions are extremely difficult to grasp. Put another way, we creatures of practical reason normally act “under the

According to source compatibilists and incompatibilists, we do not have to be able to do or act otherwise, as intuitive as this might be. Instead, we have to stand in a certain relation to our actions and to the outside world in order to be responsible for our actions. These actions have to have the right connection to us, so to speak. Think about Ann the kleptomaniac once again. Many of us believe that she is not responsible for her behavior when inside a store and shoplifting. Above, I provided one possible analysis for why this is so: she could not help herself stealing. That is, she could not have acted otherwise. Now, even if this is true, we might still think that this judgment is just “piggybacking” on something else. This conclusion has been drawn after finding cases where people seemed to be responsible even though they could not do or act otherwise, in combination with an idea about what drives our intuitions in cases such as Ann’s. I will discuss these in order.

In his seminal (1969) paper, “Alternate Possibilities and Moral Responsibility,” Harry Frankfurt presented a case meant to show that we can be substantially responsible without having the ability to do otherwise. A usual Frankfurt-style case goes something like this. White is deliberating whether he should break into a car. On the one hand, he needs a car really badly right now so that he can come home without having to walk home in the rain. On the other hand, he thinks it would be the wrong thing to do, morally, and he cares about being a decent person. After a bit of soul-searching, White decides on his own to steal the car. “It is better to be dry than decent.” Now, unbeknownst to White, Black, who is a neurology professor, has planted a chip in White’s brain that registers White’s every thought and by which Black can control White’s every move. For nefarious reasons, Black wants White to steal the car, so if White had decided not to steal a car, Black would have intervened and made him do otherwise. Put another way, White could not have done otherwise. If he had tried, Black would have stopped him. But as it happens, completely on his own, White decided to steal the car, which meant that Black could just lean back and do nothing. In these cases, many of us judge

idea of freedom” (cf. Korsgaard 1996), and it is difficult to understand how it would be to live a life under other conditions.

White to be responsible. However, White could not do or act otherwise, which means that the leeway criterion must be false or at least modified.⁶

The source explanation for why Ann is not responsible for her behavior is, instead, that she did not want to steal. It repulses her, because she is a firm believer in the Ten Commandments, one of which is, “Thou shall not steal.” Ann is not a thief; it is something outside of her that makes her do these things. Or to take another example, when people are being manipulated, we often judge them to be at least less than fully responsible for what they do. Women who are trapped in destructive relationships where they are convinced about their own worthlessness by their partners are just one example. In these cases, we might assume, these women can do or act otherwise. They can break up and leave at any point. But they are so convinced that they will not find anyone else or even manage on their own that they choose to stay. Many of us believe that the woman being manipulated in such a way is not at least fully responsible for staying, and this is intuitively explained with reference to the source view. The source of the actions performed under manipulation is not the individual who is acting but the person who is manipulating; hence, the person under manipulation is not (fully) responsible.

We have, then, different and conflicting accounts of what it takes for us to be responsible. A wide range of skeptical arguments have been launched against the compatibilist and soft incompatibilist views. The two I will discuss are that from heteronomy and that from luck, beginning with the argument from heteronomy.

3.1.2 The argument(s) from heteronomy

There is a family of arguments trying to show that we are heteronomous, that we are not in any relevant sense in control of and/or the source of our actions. These arguments from heteronomy can be teased out in a number of different ways (see for example Haji 2009). But what lies behind many of these sorts of skeptical ideas is roughly something like this. Actions are events. If the universe is deterministic, then every event is completely caused by some earlier set of events together with the laws of nature. Each of these events, in turn, is completely caused by a still earlier set of events together with the laws of na-

⁶ For a recent discussion on Frankfurt-style cases, see for example, Widerker and McKenna 2003.

ture... In principle, we could follow this causal chain back until the beginning of time. However, if the universe is fully indeterministic instead, every event just happens. The events are not causally determined and do not obey any laws in any commonsensical use of the word. Either way, whether our actions are determined or undetermined, they are ultimately caused by factors outside our control. We cannot control things that happened at the beginning of time, the laws of nature, or what just happens.

Now, many people feel, when being confronted with this sort of case, that we cannot be responsible.⁷ Depending on whether you defend a source or a leeway view, the explanations differ for our exculpatory intuitions when confronted with these cases. Begin with the leeway incompatibilists. According to them, in a universe where people's actions are ultimately and completely caused by factors outside their control, what explains our exculpating intuitions is that they *cannot avoid anything*. All actions are completely caused by previous events together with the laws of nature, and these prior events are completely caused by other prior events and the laws of nature, and so on, until we go back to the beginning of time. But, the incompatibilist says, they cannot control the laws of nature or what happened in the beginning of time,

⁷ Described in this way, many people will feel the intuitive pull toward the idea that we cannot be responsible or deserve anything. This is true for people in the health and healthcare debate as well; WHO says for example that

“...heart disease is caused not by a lack of coronary care units but by the lives people lead, which are shaped by the environments in which they live, and obesity is not caused by moral failure on the part of individuals but by the excess availability of high-fat and high-sugar foods” (WHO 2008: 35).

Here they admit that the cause of people's heart diseases is their lifestyle, but since their lifestyle is caused by something else—the availability of high-fat and high sugar foods—people are not responsible for these illnesses and hence do not deserve any sort of response because of it. Or take Alena Buyx who says that

“We justly hold people responsible for consequences of their behavior only if they had control over it and chose freely. While most health-related behaviors are not uncontrollable impulses but instead are subject to conscious choice, these choices are in turn often influenced by multiple factors: socioeconomic status, socialization and education...and so forth (Buyx 2007: 873).”

The thought here is that to the extent people's actions are determined by factors outside their control, such as socioeconomic status or socialization, people are excused for them. Since our actions at least in part are caused by ourselves, we are not fully excused. But if they were, we would be.

which means that they cannot control what they do now.⁸ Differently put, given the previous condition together with the laws of nature, the actions they perform had to happen, and nothing else could have happened instead. Their actions were inevitable. This is a rough version of one of the most famous arguments in favor of leeway incompatibilism, “the consequence argument” (van Inwagen 1983).⁹ Of course, if something in the history or the laws of nature had been different, my actions would perhaps be different as well. But in the actual situation I am in, with the laws of nature and how the actual history up until now has played out, the actions I actually perform are inevitable, and this is what matters for responsibility. What the incompatibilist wants is an unconditional (absolute) reading of “can” in “she can act otherwise,” and this reading is incompatible with our actions being ultimately determined by factors outside our control.

Soft incompatibilists believe that if our actions are determined by factors outside our control, we cannot be responsible for them. But, they continue, since it is not true that our actions are determined by factors out of our control in such a way that we cannot do anything other than what we did, the argument from heteronomy does not preclude responsibility. Some argue that actions are not events but are instead *causa sui* and that human agents are primary movers who are unmoved (Chisholm 2003). These accounts have been criticized on the grounds that they are mysterious and have no empirical evidence in their favor (Pereboom 2001, Persson 2005). However, in recent years, many clear and naturalistic soft incompatibilist accounts have been put forward (Kane 1996, Clarke 2003, Mele 2008, Balaguer 2010). Needless to say, these solutions are extremely controversial as well (Strawson 2001, Persson 2005, Pereboom *forthcoming*), but I will not discuss this further here.

The leeway compatibilists often agree that there is something to the arguments from heteronomy, such as the consequence argument. However, they reject the incompatibilist idea about the absolute reading of “can,” arguing that we only need a conditional “can” to be responsible. For example, what

⁸ This is true, at least, for finite beings such as humans.

⁹ It should be noted that “determinism” does not necessarily do any work here. If my actions instead of being determined just happened, this would not give me any more control, i.e., any option to avoid doing what I did.

people need in order to be responsible is that they could do something else *if they wanted to*.¹⁰ Ann is impaired in this way. She cannot do what she wants to do, and this is why she is not responsible. The problem is not that her actions are ultimately determined by factors outside her control. Rather, it is a much more mundane personality and/or brain dysfunction of some sort that diminishes her ability to do otherwise, compatible with her actions being determined by factors outside her control.

In addition, compatibilists often try to explain away the seeming force of the consequence argument and other arguments from heteronomy, arguing, for instance, that we easily mix up fatalism and heteronomy (Dennett 1984: 128–130). In a fatalistic scenario, no matter what you do, whatever was fatalistically determined to happen will happen. If this were true with respect to the scenario described in the consequence argument, we would not, according to the compatibilist, be responsible. But this is not true. If we wanted to do something else, we did it, and our actions were causally effective, something else would happen. According to compatibilists, this suggests that the consequence argument is false. The upshot is that if this compatibilist line of reasoning were to succeed, we would need a less demanding argument to show that people can be responsible. After all, most people would agree that many of us in numerous situations have these conditional abilities. Of course, incompatibilists do not agree about this conditional reading of “can,” and discussions go back and forth, but according to many philosophers, the discussion about the right reading of “can” has reached a stalemate (Fischer 1994, Kane 1996). This means that it would be a good thing if we could find a new angle in discussing this question.

As we have seen, there are both source compatibilists and source incompatibilists. Source compatibilists believe that the agent and her actions have to be related in certain specified ways for her to be responsible for them. For example, a child who is brought up not to think for herself, through constant physical abuse, might not be responsible for what she does as an adult when she acts on these “unsheddable” dispositions, i.e., dispositions she cannot get rid of. This, according to one source compatibilist account, is because she is not

¹⁰ There is a vast discussion about how these conditionals should be spelled out. See Kane 1996: 44–59 for references.

the source of her actions; her parents are (Haji and Cuypres 2004, 2007). But since this source view is compatibilist, it implies that being responsible is compatible with the actions being determined by factors outside her control. Incompatibilists believe that this is false. For us to be the source and hence be responsible for our actions, these sources ultimately have to be under our control and/or ultimately reside in our agency. Current discussion about *what sort of source* we have to be, except for relying on formalizations of different forms of heteronomy arguments, revolve around cases with agents whose actions are caused by covert manipulation (Kane 1996, Pereboom 2001, *forthcoming*). Many of us assume that we are not responsible for actions caused by manipulation. What source incompatibilists argue is that there is no relevant difference between actions caused by covert manipulation and actions caused by factors outside the agent's control. Compatibilists often agree that manipulation undermines responsibility and hence, also agree that there must be a relevant difference between the two.¹¹

Probably the most widely discussed manipulation argument is Derk Pereboom's "four-case argument" (Pereboom 2001: 112–116).¹² In each of Pereboom's four hypothetical cases, Professor Plum murders his colleague, White, to get an advantage in his career. In the first case, Plum is constructed and directly controlled by a group of neuroscientists who cause him to think and react in a way that makes Plum kill White. In the second case, the neuroscientists do not control Plum directly, but they have constructed him in such a way that when he finds himself in these sorts of situations, he thinks and reacts in a certain way, and this causes him to kill White. In the third case, the neuroscientists have been removed. Instead, Plum comes from a community where he has been rigorously trained to react in certain ways in these circumstances, which causes him to kill White. In the last case, Plum is living in a world in which his actions are caused by factors outside his control, where no training is taking place and where, under the same conditions, he kills White.

¹¹ But see McKenna (2008) for another view.

¹² This is in competition with the "Ann/Beth" and "the zygote" argument posed by Alfred Mele (1996, 2006). For a discussion about the four-case argument, see for example McKenna 2004, 2008, Mele 2005, 2006, 2008, Pereboom 1995, 2008, *forthcoming*.

Presumably, most of us would say that Plum is not responsible in the first and perhaps not in the second case. However, most would judge him responsible in the third, depending on how we work out the details of the case, and at least compatibilists would want to say that Plum is responsible in the last case. Pereboom's challenge for the compatibilist is to find the relevant difference, between the first and second cases and between the third and fourth cases, that explains why Plum is responsible in the latter but not in the former cases. This is of course quite a simple thing to do with my sketchy presentation, but harder when details are filled in. For example, one might think that what is doing the work here is the fact that agents are controlling other agents. But if we replace the agent with spontaneous computers that in some random way just started to exist completely out of the blue, and if they in turn started to create and control people in the same way that the neurologists did, we would still believe that people like Plum, who are under the influence of these computers, are excused (Pereboom 2001: 116).

The discussion about manipulation has vitalized the debate on conditions for responsibility, and many new and important distinctions have been made.¹³ For one, the difference between the source and leeway views, irrespective of whether we are discussing compatibilism or incompatibilism, has become much clearer. For another, new distinctions have been made within compatibilism and incompatibilism. For example, source compatibilists are today defining themselves as historical or nonhistorical, where the historical compatibilists believe that the history of how the agent came to be is relevant to whether she is responsible (Fischer and Ravizza 1998, Mele 1996, 2006) while the nonhistorical (time-slice) compatibilists do not (Dennett 2003, McKenna 2008); this is something quite new in the debate. This distinction is

¹³ Another intriguing case against source incompatibilism comes from Mikael McKenna (2004), about a "magical agent" called "Suzy instant." Suzy is (biologically) a normal adult around "thirty something," with all the cognitive abilities and knowledge in place to make fully rational choices. The difference between Suzy and other women the same age and in the same circumstances is that Suzy was just spontaneously created out of thin air. The "magic" here is that this does not normally happen, but in this case, it has. So it is not the case that someone created her with some purpose in mind; her existence is just a cosmic coincidence. Now, just after she had been created, she is thinking about whether to give some of her money to a beggar on the street. Because she is a good person who thinks it is important to give to the less fortunate, she reaches into her pocket and gives some money away. The question here is whether or not she is responsible. Many people would say yes, even though she is not responsible "all the way down" as the source incompatibilists contend.

especially in the discussion about responsibility and health, since many people believe that the effects of advertising make people less responsible for acting in an imprudent way (Abraham 1993, Atterbury 1996, Übel et al. 2001). It is well established that advertising has effects on alcohol consumption, use of cigarettes, and eating of unhealthy foods. Even if this is true, according to some nonhistorical views, people are still fully responsible for their behavior given that they wholeheartedly endorse it.

With all this being said, it seems to me that the debate has become much less innovative in recent years. People have dug in their heels and are defending the positions they hold, arguing that their compatibilist or incompatibilist theories can explain the intuitions at hand, and there are now fewer steps forward being made. So even though this debate has been invigorating, it seems that it has now come to somewhat of a halt.

3.1.3 The argument(s) from contrastive luck

I mentioned two kinds of skeptical arguments. The one we have discussed up until now was heteronomy. The other skeptical argument is the one from moral or contrastive luck (see for example Finberg 1970, Nagel 1979, Williams 1981, Kane 1999, Clarke 2005, Mele 2006).¹⁴ I will mainly use the term “contrastive” here, since there might be much luck involved in cases where responsibility is undermined but where it is not clear that anything of moral value is at stake.

Most people believe that we have to be in control in order to be responsible (i.e., the control condition). Some think that what is required is ultimate control, being the ultimate source of one’s actions, while others think that it requires alternate possibilities, that is, the ability to do or act otherwise. Many also believe that we should judge two people as being responsible to the same extent if the only difference between them is beyond their control. This latter claim leads to a range of different forms of problems for different responsibility theories, because when you compare people in this way, it seems that there is always much luck involved, in the sense that, for instance, identical actions performed by different people lead to radically different outcomes because of factors outside these individuals’ control.

¹⁴ For an excellent overview of the debate, see Nelkin 2013.

According to Thomas Nagel (1979), there are four types of contrastive luck: resultant, circumstantial, constitutive, and causal. *Resultant* bad luck is when two individuals do the same thing but only one suffers from a terrible outcome because of factors outside her control, as when two people who live in different places smoke and only one is afflicted with lung cancer. Because of indeterministic processes in the cell leading up to damage on the DNA, the one who is afflicted suffers from bad resultant luck. After all, assuming that it was not possible to know about these indeterministic processes, the difference between the two individuals, that one got cancer while the other did not, was outside the individuals' control, and hence, perhaps something that they were not responsible for.

Circumstantial bad luck is when two individuals would do the same thing if the circumstances were the same, but since the circumstances differ and the individuals have no control over them, one person does something that has a terrible outcome while the other one does not. For example, if we have two individuals equally inclined to become alcoholics, but alcohol is banned and non-existent in the country where one person lives, while being socially respectable and widely available in the country where the other person lives, then it is just a matter of circumstantial luck that the latter person drinks alcohol and develops a severe alcohol addiction while the former does not. Assuming that the addicted individual is not in control of where she was brought up and lives, it might be the case that she is not responsible for her addiction either.

Constitutive bad luck is when two individuals are disposed in different ways, in the sense that they have different traits and dispositions because of factors outside their control, where this leads to a terrible outcome for one but not for the other. For example, two people are living in the same alcohol-soaked culture, but one is an alcoholic while the other one is not. Assume that the only difference between the two is that the alcoholic has a strong genetic disposition to be an alcoholic while the nonalcoholic has not. Supposedly, we have no control over genes; hence, the difference between the individuals is a matter of luck. This might mean that the alcoholic is not responsible for being an addict rather than not.

The last category of contrastive luck is that of *causal* bad luck. This is simply making the argument from heteronomy contrastive. Again, take two individuals who desire different things and so do different things. In a deterministic scenario, our desires are ultimately determined by factors outside our control, which means that the reasons why one person has a will that leads to worse results than the other is not something under his control and is also, perhaps, something for which he is not responsible. And as we saw above, something similar can be said about the indeterministic scenario. There, things just happen, and the fact that things just happen does not increase our control; hence, we are not more responsible in that scenario than in the deterministic one.¹⁵

For compatibilists and soft incompatibilists, there are two main strategies for handling these questions about contrastive luck: either they hold on to the control condition and try to deny the existence of moral luck through making the discussion more precise (Thomson 1993, Wolf 2001, Zimmerman 2002), or they give up the control condition proposed above and argue that we do

¹⁵ It is important to note that the discussion about “resultant luck” has also been cast in terms of “causal responsibility.” For example, it is commonly held in the debate on responsibility for events in general (Driver 2007, Vallentyne 2008, Temkin 2011) and for responsibility for health in particular (Olsen 1997, Wikler 2004, Daniels 2008) that having caused these events is a necessary condition to be responsible for them.

Many people in the debate on causal responsibility believe that in a case like that with the two individuals who smoke to the same extent but are afflicted differently, it follows that these individuals are not fully causally responsible. Larry Temkin says, for instance, when he discusses causal responsibility, that “Most people who overeat don’t have a stroke, and most helmetless motorcyclists don’t end up in the emergency room. Thus...one must remember that full responsibility for one’s choices doesn’t entail full responsibility for one’s predicament” (Temkin 2011: 65). His idea here is that if *every* motorcyclist and overeating individual were afflicted with injuries and/or ill health, they would be fully causally responsible and perhaps also fully substantially responsible. But since not everyone is afflicted, they are not fully causally responsible and hence are not fully substantially responsible either. Similarly, according to Peter Vallentyne, “[A]n agent is responsible for an outcome to the extent that her choice increased the chance of the outcome and she believed this to be so” (Vallentyne 2008: 64). (More precisely, Vallentyne argues that “...an agent’s choice, *c*, increases the chance of outcome, *o*, just in case, just prior to the choice (given the state of the world and the laws of nature), the objective chance of *o* given *c* is higher than the objective chance of *o* given that *she* makes an autonomous choice (of some sort)” (Vallentyne 2008: 63)). The increased chance he is talking about here is objective in nature, and it is based on single case propensities (Vallentyne 2008: 63).

The idea, then, is that given the history of the universe and the laws of nature, my decision and in the end my action increased the chance that an event would happen. This explains why the motorcyclist was more responsible in the case when the risks were great than when the risks were low. Her choice increased the risks of the event happening in the former case in comparison with the latter case.

not need to fulfill it in order to be responsible (Adams 1985, Walker 1991, Moore 1997, 2009).¹⁶ I will here give one example of each so that the picture becomes a bit clearer.

Michael Zimmerman argues that we have to distinguish *degree* from *scope* when discussing resulting contrastive luck (Zimmerman 2002: 560--562). His example is that of George who shoots Henrik with his rifle and where Henrik dies in one version of the case and survives in the other version.¹⁷ What Zimmerman argues is that Henrik is equally responsible in both cases, i.e., the degree of responsibility is the same, and that he in both cases deserves the same treatment.¹⁸ The difference is the scope of what George is responsible for. George is responsible for more things when he tries to and actually kills George than when he only tries to kill George. Now, it might be thought to be a bit strange that if you actually are “responsible for more things” and these things are quite bad things (such as an individual dying), you are not more blameworthy than when these “things” are not as bad (such as an individual being severely injured) and do not deserve a harsher treatment. But Zimmerman has defended a theory on responsibility which, he argues, can account for these intricacies (Zimmerman 1987). What you can say is that there are many notions of responsibility, and the notion employed here by Zimmerman is not what I have called “substantial responsibility,” which means that Zimmerman’s solution is not a solution viable for us.

An example of someone who has discussed the problems of contrastive luck from the other side, so to speak, is Michael Moore, who argues that we need to change our conception of control to account for these problems. What he does is to begin, in particular judgments about people’s characters and actions, by showing that in these ordinary cases, we deny the relevance of the “contrastive control condition.” For example, we feel greater guilt when we actually succeed in doing something terrible than when we just try and fail (Moore 1998: 230–232). He then he goes on arguing about how we philoso-

¹⁶ Of course, some would want us to give up the notion of responsibility altogether because of contrastive luck (e.g., Levy 2011).

¹⁷ This is not entirely true; Zimmer uses four different characters, Henry/Henrik and George/Georg, but it seems to me that this simplification is not distorting the main point.

¹⁸ Needless to say, there might be other reasons for treating people who actually succeed in their pursuit of murder than those who fail.

phers need to change our conception of control, and why it seemed to be a viable principle in the first place.¹⁹

Moore believes that there are relevant forms of “moral luck,” which is the occurrence of effects created under freakish circumstances way out of the ordinary (Moore 1998: 214–217). If I aim and shoot my rifle at a moose, and my shot hits another hunter next to the moose because an abnormally strong gust of wind took grip of the bullet, I might not be responsible for hitting my fellow hunter. But the ordinary cases of moral luck are not like these cases. People who aim at and shoot others do not hit and kill them because of freakish accidents. Instead, it is quite a commonplace, at least if you are decent marksmen, are in a reasonable range to your target, and the weather conditions are OK, to hit and kill the individual you shoot at. This implies that these people who shoot and hit should not be excused on the grounds of moral luck.

Of course, there is a lot to say about Moore’s and other similar arguments. I will not do that here. What we can bring with us is that there will still be plenty of cases where it seems intuitive to many people, such as Zimmerman, that contrastive luck in different forms when described in certain ways, elicit skeptical intuitions regarding our ability to be substantially responsible. To disregard these cases will be continuously controversial, especially since not everyone will agree that the debunking attempts, such as those we have seen being deployed here by Moore, work.

Connecting the discussion of contrastive luck more tightly to the discussion about the different theories on responsibility in section 3.1.1 above, we find that depending on the view of what is required for responsibility, different answers to this challenge will be forthcoming. Begin with the soft incompatibilists. Some of them argue that their theories are not afflicted with some of the variants of contrastive luck such as constitutive and causal, because we are ultimately responsible for our actions and hence are the ultimate causes of our own wills, our character, and hence what we do in these different situa-

¹⁹ The same strategy has been used in the discussion of culpable ignorance (see for example Sher 2009). Sher argues that the “searchlight view,” the view that we need to fulfill quite extensive epistemological requirements to be responsible, only seems to be intuitive because it is piggy-backing on another more reasonable view that has to do with the notion that the motivational structure of the individual has to be able to explain the outcome for which the individual is responsible.

tions. Soft incompatibilists often readily concede to the fact that we might not be fully responsible for the results of our actions, but they still believe us to be responsible for much else.²⁰

Even though for the soft incompatibilist answer too, many forms of contrastive luck might appear convincing on first glance, the problem of contrastive luck still arises for those who insist that at least some of our choices must be undetermined up until the moment they have been made.²¹ This is true for both soft leeway and source incompatibilists. For example, the source incompatibilist Robert Kane (1996) believes that some decisions have to be undetermined in order for us to create ourselves and hence to be responsible for what we do. Others, who believe that we do not have to create ourselves in this way to be substantially responsible, still believe that at least some decisions must be undetermined in this way for them not to be determined by factors outside our control (O'Connor 2000, Clarke 2003).

The problem for soft incompatibilists is that to have genuine indeterminacy at the moment of choice, two individuals who are identical in every respect, living in worlds identical in every respect, and finding themselves in identical situations could end up doing radically different things without anything accounting for this difference in choices (Mele 2006). Accordingly, the difference seems to be a result of luck, which in turn might relieve us from our responsibility. The discussion revolves around cases where we are invited to compare two identical worlds, W and W' , where John, in the actual world W , decides to kill Bill, and in the possible world W' , John decides not to. Given that there is no difference between W and W' up to the killing, it seems that John cannot be responsible for “killing Bill instead of not killing Bill,” and then we might wonder whether John actually is responsible for killing Bill. After all, there is nothing about John that can explain why he killed Bill in W but not in W' . And if there is nothing to explain this, it just seems to be a fluke, something that one would generally not be considered responsible for.

One general strategy in trying to answer this luck objection has been to be more precise about what we are responsible for and to supply some cases

²⁰ See for example Vallentyne 2008.

²¹ It has been argued that this is a problem for compatibilists as well (Björnsson and Persson 2012a, Levy 2011: 84–110).

where we intuitively seem to be responsible but where indeterminacy is involved. For example, it is said that John is responsible for killing Bill, and this becomes intuitive if we focus on the actual scenario and what led up to it instead of on some sort of contrastive scenario (see for example Kane 1999, Clarke 2005). What is relevant, then, is what John's reasons actually were, whether he actually planned for this to happen, and so on. Of course, the question then becomes why we should focus on the actual scenario and not the contrastive scenario, and here the discussion has reached somewhat of an impasse.²²

Compatibilists, on the other hand, might seem to be in the clear. They defend accounts where people are responsible even though their actions are determined by factors outside their control. Certainly many of them agree that some sorts of luck might mitigate responsibility. But they still think that many of the other problems might be solved through, for instance, being clear on what we are responsible for and which factors are exculpating. For example, if someone came up and bashed our head in so that we suffered a massive head injury and, because of that, lost most of our ability to be rational and govern our impulses, then we are excused (cf. Dennett 2003). But this is completely different from the cases of contrastive luck. Without going into the nitty-gritty of it all, it seems that those who agree that some contrastive luck undermines responsibility must also defend why these sorts of luck and no other forms undermine responsibility.²³

²² This is similar to the discussion about causal responsibility. Traditionally, in tort law, the so-called "but-for" test, the test for strong necessity, has been applied when trying to find out whether a defendant's act has caused a certain consequence (Honoré 1995). Take a case involving smoking. If we assume that were it not *but for* the person smoking, she would not suffer from lung cancer, it means that her actions, according to the but-for test, were a cause of her getting lung cancer, and this in turn explains why we, at least today when most people know of the risks, deem people to be responsible for getting lung cancer when they smoke.

However, one problem with these accounts is that of overdetermination. For example, Sam kills the mayor, but unbeknownst to him, Jack would have killed the mayor if Sam did not. In these cases, many people think that Sam is responsible and that his actions were the cause of the mayor's death, even though the mayor's death would not have happened but for Sam's actions. If Sam had not killed the mayor, the mayor would still have been killed by Jack. Those in favor of the but-for account have solved the puzzle in a way similar to the incompatibilists above.

²³ For an account of why (historical) compatibilists have similar problems to soft incompatibilists, see Levy 2011: 84–110.

Furthermore, they also have to convince others that the exculpating intuitions they experience are flawed somehow. Go back to the question on causal contrastive luck: assume that there is no difference between John who killed Bill in world W , and John in world W' who did not kill Bill. Instead, something else on the outside of John, not under his control, caused John to kill Bill in W but not in W' , like the warm sun and an old tree branch falling to the ground. (Think, for example, about *The Stranger* by Albert Camus). Or perhaps something was different on the inside of John but not under his control, such as neurological processes subtly moving him over the brink to murder. Whatever the differences may be, we can assume that it is nothing about John that made the killing happen, in the sense that he had different beliefs, desires, experiences, and so on. It is just that he was unlucky when killing Bill instead of not killing Bill. Since many people have this intuition about causal contrastive luck, this needs to be explained away. And even though this has been tried, no consensus has been reached.

3.1.4 Summary and outlook

In conclusion, we have a range of different positions and problems with regard to the conditions under which we are responsible. Both the cases of heteronomy and luck imbue many with excusing intuitions. However, there is no consensus on how we should solve these problems and bring our intuitions in line. Instead, different stalemates have been reached and heels have firmly been dug in, which means that a new perspective on the debate is in order.

PAPER 1: THE EXPLANATORY COMPONENT OF MORAL RESPONSIBILITY

In section 3.1, we saw that there are many views about the conditions under which we are responsible and how these views are related to the different skeptical arguments. In this paper, Gunnar Björnsson and I address question 1 by explaining why we react the way we do when confronting these different cases and what implications this has for the debate on the conditions for responsibility. We do this by devising a psychological model, the *Explanation Hypothesis*, combining the obvious idea that there is some sort of explanatory relation imbued in the notion of responsibility and the commonsensical observation that explanatory judgments are sensitive to expectations and interests in different ways. This leads to new and interesting results.

According to the *Explanation Hypothesis*, we judge people to be responsible when we perceive their motivational structure to have a *general tendency* (General explanatory tendency (GET)) to be a *significant explanation* (Explanation Responsibility (ER)) of a *relevant kind* (Reactive Response-ability (RR)). Applying the explanation hypothesis, we can account for our responsibility judgments in the context of ordinary excuses and with regard to the skeptical arguments of heteronomy and luck.

According to ER, when we judge someone to be responsible for a certain event, we judge her motivational structure to be a significant explanation for that event. However, what we judge to be a significant explanation has to do with our expectations and explanatory interests, that is, our explanatory frame (explanatory perspective). There are many causally relevant and necessary factors required for the event to happen; the reason why we pick one or some of these as significant is that they are unexpected somehow. For example, when the fire alarm goes off and there is smoke in the air, we usually cite the smoke as the explanation of why the alarm went off instead of the fact that the alarm is in working order, that it was turned on, etcetera—factors that were causally necessary for bringing about the alarm's going off. The reason for this is that we do not usually expect smoke, but we do expect the alarm to be in working order. Now, if instead we had expected the fire alarm to be turned off, for instance, because we knew that smoke was going to be present, then we would cite the alarm not being turned off as the explanation for why it went off.

The same thing can be said about normative expectations. For example, assuming that a police officer could prevent a robbery if she only did her job instead of listening to the dog races, the explanation of why the robbery was successful would be that the police officer was more interested in dog races than the street life. The reason why we judge the police officer to be responsible is that we have normative expectations of people doing their jobs, especially when doing otherwise would put others at great risk. The reason, then, why we cite the officer's inattentiveness instead of the robber's skill, say, is that the officer's motivational structure stands out because of her normative transgression. Of course, many people would also cite the robber as responsible for the outcome as well. This is what to expect, according to EH, because he commits an act perceived as grossly immoral. However, judging two individuals as responsible for the same event, such as the robber and the police officer, is often followed with a certain amount of tension. ER can account for this as well: when judging the police officer as responsible, we take an explanatory frame where the robber's motivational structure falls into the explanatory background and vice versa; hence, when perceiving one motivational structure as significant, the other becomes insignificant. We can take a step back and widen the explanatory frame so that the conjunction of both actions is perceived as an explanation to the event. But this is a more abstract take on the matter, and hence more cognitively complex, which makes the tension not go away completely.

The *Explanation Hypothesis* can also account for the persuasiveness of skeptical arguments from heteronomy and contrastive luck. It is easy to see that the persuasiveness of these arguments has to do with the fact that the motivational structure falls into the explanatory background and hence becomes an insignificant explanation.

As we saw in section 3.1, the arguments from heteronomy introduce an abstract way of perceiving people's actions, decisions, etc., so that all explanatory significance of these structures disappears. The arguments are without any detail about the motivational structure that could serve as an explanation, and previous causes are explicitly mentioned as possible explainers, which mean that it becomes much easier to follow the causal chain backward until we reach a state where our motivational structure either does not exist or where it

does not fulfill RR. Also, when factors such as heredity and early experiences are explicitly mentioned as previous causes, we start to expect that motivational structures of individuals can be straightforwardly explained by something else. When the heredity and early experiences come separately or together, many of us judge the motivational structures of the individuals not to be significant explanations of anything. Because we see them as mere links in the causal chains leading to be individuals' actions, we do not see the individuals as responsible. If the argument were to be described in more concrete terms, explicitly mentioning the individuals' actions, beliefs, desires, etcetera, these heteronomy arguments would not be as persuasive (more on this in Paper 2 and in section 3.2, below).

Similarly, in skeptical arguments from contrastive luck, the motivational structure of the individual falls into the explanatory background. The discussion revolves around cases where we are invited to compare two identical worlds, the actual world W and a possible world W' . In the actual world, John decides to kill Bill, whereas in the possible world W' he decides not to. Given that there is no difference between John in W and W' up to the killing, it seems that killing Bill is just a matter of "luck" and hence something that John is not responsible for. As we saw in 3.1, the soft incompatibilists' comeback was that we should focus on the fact that we judge John to be responsible for *killing Bill* and not on what we might judge when focusing on what John would have done in another possible world.

Again, the explanation hypothesis can easily account for why both these sets of judgments are made. When we compare both worlds with two different outcomes, and we want to explain why these two different outcomes have come about, we need to find some difference between these two worlds. In the cases of contrastive luck, no such difference is to be found, which means that we cannot find anything that could explain the difference in actions. It seems just to be a stroke of bad luck that John killed Bill instead of not killing him, and with no explanation at hand, the explanation hypothesis implies that we get the intuition that John is not responsible for doing instead of not doing. This exculpatory intuition is the one that many people get when confronting these cases of contrastive luck. On the other hand, when we focus the way the soft incompatibilist wants us to, we will be more inclined to judge

John to be responsible. Then we can more clearly see how his motivational structure violates our normative expectations and hence becomes a significant explanation for why the murder took place.

In the bulk of the paper, we argue that the explanation hypothesis can account for our intuitions when confronting different skeptical arguments, because these cases make us adopt different sorts of explanatory frames. The implication, we furthermore suggest, is that the skeptical arguments lose their evidential value. We know why these arguments elicit strong intuitions: they make us adopt an explanatory perspective where people's motivational structures do not seem to significantly explain the outcomes at hand. However, there are other non-skeptical perspectives to take; hence we need some argument in favor of picking one perspective over the other. We suggest instead that when trying to answer the question on what perspective we should take, there are two routes to take: one metaethical (semantic) and one normative. What does it *mean* when we say that someone is responsible (the semantic route), and what *should* we care about when judging someone to be responsible (the normative route)? In the last section of the paper, we argue for the idea that some of the most prevalent semantic theories (contextualism, realism, and expressivism) and normative theories (consequentialism, contractual theories, and virtue ethics) are all in favor of taking an everyday non-skeptical perspective.

3.2 RESPONSIBILITY AND EXPERIMENTAL PHILOSOPHY

One of the overarching questions for this dissertation is whether we are responsible for our own ill health. To answer this question, we have to know under which conditions we are responsible, a question that we have seen is difficult to answer. There is a wide range of possible positions to take, and many stalemates have been reached. In recent years, some philosophers have turned to “experimental philosophy” in their pursuit of loosening up these locked positions. In this section, I will first discuss what experimental philosophy is and how it has developed as a philosophical field.²⁴ Thereafter I will discuss its relevance for our pursuit of finding a theory for responsibility judgments, before I say something about the accounts trying to explain the results found by experimental philosophers and why they fail. In the end, I will give a summary of my second paper, where a unified explanation of responsibility judgments is presented.

3.2.1 What is experimental philosophy?

In experimental philosophy, you mainly study the intuitions of non-philosophers, “the folk.”²⁵ One aim is to explore whether philosophers’ intuitions are shared by non-philosophers and what might cause these intuitions. In the basic form, the experimental philosopher presents people with thought experiments to see and measure how they react to them. Take the case of free will and determinism.²⁶ Many philosophers think that incompatibilism, the notion that determinism and free will are incompatible, is the natural position to take in a discussion about free will. This idea about incompatibilism being natural has implications for the debate. For example, when Kane discusses the issue of compatibilism (the idea that determinism is compatible with free will) versus incompatibilism, he says:

²⁴ For a nice introduction to experimental philosophy, see Alexander 2012.

²⁵ However, as we will see in section 3.2.2, these studies now include professional philosophers as well.

²⁶ Even though free will and responsibility are related, there are those who believe that, for instance, determinism and responsibility are compatible but not free will and determinism (e.g., Fischer and Ravizza 1998). See also footnote 4.

...is compatibilism believable? In my experience, most persons resist the idea that free will and determinism might be compatible when they encounter it...If compatibilism is to be taken seriously by ordinary persons, they have to be talked out of this natural belief in incompatibility of free will and determinism by means of philosophical arguments (Kane 2005: 12f).

Kane then goes on to paint a picture of how the compatibilist persuades the layperson to believe that her position is correct. Thinking that incompatibilism is the natural position, many philosophers have thought that it is compatibilists who have to supply the arguments for their position while the incompatibilist can sit back and try to show why these arguments are false. But is incompatibilism really the natural position? What do people think? A range of studies have found that laypeople have *both* strong compatibilist and strong incompatibilist intuitions (see for example Nichols and Knobe 2008, Nahmias et al. 2009). What people think depends, for example, on how cases are described or framed, to what extent we get to know anything about the people in the cases, and so on. The point is that incompatibilism should not be taken as natural without substantial argument. Since this argument is not yet presented, it should change the dynamic of the debate and perhaps loosen up some stalemates.

Another aim of experimental philosophy is to find the underlying psychological mechanisms which explain what drives our intuitions when being confronted with different cases and arguments. Some of these findings have moved people to switch positions in different debates. As an example, Nozick's (1974) classic case against hedonism asks whether a pleasurable life lived in an experience machine is as good as an equally pleasurable life outside the machine, where all the experiences you have actually happen for real. The common intuition among philosophers, including Nozick himself, is that the life outside is better in comparison with the one inside the machine. When probing laypeople's intuitions in experiments, it has been shown, however, that peoples' answers differ depending on whether they are asked if they want to hook up to the machine or whether they want to *continue* being hooked up to the machine (e.g., de Brigard 2010). If subjects are asked to imagine that they are already hooked up, they want to continue to be so, and when they imagine

not being hooked up, they want to continue to be so as well. Accordingly, the underlying mechanism that explains why we believe the life in the experience machine is less valuable seems to be conservatism and not that there is a certain value in having real and meaningful relations to actually make certain accomplishments and so on. Because of this finding, among others, the well-renowned philosopher Peter Singer, who has been a preferentialist for his entire career, has moved to become a hedonist.²⁷ This suggests that results from experimental philosophy can change the minds of philosophers and perhaps can also refocus the philosophical debate.

However, there has been much debate on the actual use of folk intuitions in philosophy and critique against the experimental philosophy movement (see for example Kauppinen 2007, Kappelen 2012). There are both concerns regarding the methods used and the ambitions of the project. The most common critique is that experimental philosophers only reach “surface intuitions” that are heavily distorted by irrelevant aspects of the cases and arguments under scrutiny. Philosophers, on the other hand, are able to transcend these different sorts of biases, reaching stable positions, through exposing themselves to a wide range of arguments.

Of course, there is a lot to say about these views.²⁸ Certainly there might be a good case against experimental philosophers in that they have not yet found what the folks’ robust intuitions and the underlying mechanisms are. But with increasingly larger groups of people being tested with a greater range of vignettes that are analyzed by ever more powerful statistical measures, it is likely that experimental philosophers will at least come closer to their goal of analyzing what the folk actually intuit.

Furthermore, it seems a bit unclear, why philosophers would be better suited to avoid reacting to irrelevant aspects in the cases and arguments. For example, think about someone who is interested in the question about the good life and who comes across Nozick’s experience machine early in her career. She is persuaded by it because she is presented with the version where we have to choose whether or not we want to hook ourselves up, and she refuses because

²⁷ He mentioned this in the Foerster Lecture Series at Berkeley 2012.

²⁸ For a to my mind convincing response to the common critique of experimental philosophy, see Nadelhoffer and Nahimas 2007.

she is affected by the very common “conservative bias.” Her belief gets stronger as she goes along in her career, partly because of other biases such as the “confirmation bias,” which makes her find a range of new evidence in favor of the falseness of hedonism, and the “disconfirmation bias,” which makes her overly critical of arguments in favor of hedonism. In the end, she concludes that pure hedonism is false. However, this conclusion has been reached, at least partly, by what many people would say is faulty reasoning and reaction to irrelevant aspects of the different arguments and cases. The point is not that philosophers are more prone to make mistakes; rather, the point is that it is difficult to see why philosophers would be immune to general psychological mechanisms and hence be in a better epistemic position than the folk, especially if the philosophers are unaware of why they react as they do when confronting different cases.²⁹

3.2.2 Folk intuitions and responsibility judgments

When discussing conditions for responsibility in the previous section, we saw that the discussion was extensive, with a variety of positions. Experimental philosophers have tried to discern which, if any, of these positions actually is the natural one and what the underlying mechanisms driving our intuitions could be. In the following, I will examine whether the skeptical arguments, those from heteronomy and contrastive luck, get support from the folk and, if so, why. First, I will briefly discuss research on how we perceive our agency.

According to many incompatibilists, their accounts of free will are supported by the way people generally perceive their agency: we feel that the future is open in an incompatibilist sense, what we are going to do is not yet determined, and at the moment when we act, we could have done something else

²⁹ This point has been substantiated by a recent study by Schwitzgebel and Cushman (2012) that shows that professional philosophers are not immune to framing biases even with regard to familiar cases and principles. Schwitzgebel and Cushman argue that philosophers’ judgments should not vary depending on the order of how different cases are presented; they should be stable. Therefore, they tested for the stability in moral judgments of professional philosophers. For example, as a professional philosopher, you should not be more inclined to accept the doctrine of double effect just because you were presented moral luck cases before. However, Schwitzgebel and Cushman found that professional philosophers are heavily influenced by these ordering effects (Schwitzgebel and Cushman 2012: 141–147). They are, for example, substantially more inclined to endorse the doctrine of double effect after being presented with the case of moral luck. This suggests that philosophers’ judgments are not as stable as many of them believe with regard to their moral thinking.

even though everything up until the act had stayed the same (van Inwagen 1983, O'Connor 1995). Of course, this does not necessarily imply that this sort of free will is needed for us to be responsible, but since this purported phenomenology lends credibility to these incompatibilist views, it might be interesting to see if they actually have this support among the folk. However, not many studies have been made on the topic, which is something to remedy but also something that makes weaker the position that the incompatibilist view on agency is the natural view. At least two studies, one by Nahmias and colleagues and one by Deery and colleagues, has, however, produced different results (Nahmias et al. 2004, Deery et al. *forthcoming*). Nahmias and colleagues found that the folk are compatibilist and Deery and colleagues that the folk are incompatibilist. As Deery and colleagues tried to amend different problems with the Nahmias study, I will concentrate on the results from Deery and colleagues.

Deery and colleagues performed three studies in which they presented vignettes to subjects and measured their reactions. The vignettes addressed a range of different topics, but all of them were posed from a first-person perspective, e.g., “You are sliding down a mountain slope...,” “When deciding what option to choose,” in order to capture the phenomenology of choice. If the subjects answered that they had the ability to do otherwise, they went through a “training section” that explained determinism, followed by control questions to secure that they had understood. Thereafter, the subjects answered questions about whether determinism was compatible with their being able to do otherwise. The results were that incompatibilistic people, who thought that they were able to do otherwise, interpreted this in a way incompatible with determinism (Deery et al. *forthcoming*: 10–31).³⁰ In general, the scores were about 5.5 on a 1 to 7 scale where 7 was complete agreement with an incompatibilist statement.³¹

³⁰ My reference is to the page number of a draft from July 29, 2012.

³¹ Shaun Nichols (2004) found something similar. It seems that young children are “agent-causalists;” they believe that we are the ultimate sources of our actions and that we can act otherwise in a strong non-conditional sense. He asked children about whether people “had to” do different things such as steal or buy ice cream, if the conditions were exactly the same (Nichols 2004: 487–489). The children stated that no, they did not have to do these things; hence, according to Nichols, we have reason to believe that children are natural incompatibilists.

If Deery and colleagues are right, it seems that we have some folk who support the incompatibilist notion of agency, at least, when they are in conditions posed from a first-person perspective. Interestingly enough, however, this does not seem to hold when people are confronted with conditions where they take a third-person perspective. In a recent study, Murray and Nahmias (2012) tested the consequence argument, i.e., the argument that if we are determined by factors outside our control, we cannot do or act otherwise in an absolute incompatibilist sense, which, in turn, suggests that we have no free will in any robust sense and hence cannot be responsible. When Murray and Nahmias tried to find out whether the consequence argument is intuitive to the folk, they let the subjects read the following vignette:

Imagine there is a universe (Universe C) that is re-created over and over again, starting from the exact same initial conditions and with all the same laws of nature. In this universe the same initial conditions and the same laws of nature cause the exact same events for the entire history of the universe, so that every single time the universe is re-created, everything must happen the exact same way. For instance, in this universe whenever a person decides to do something, every time the universe is re-created, that person decides to do the same thing at that time and then does it. (Murray and Nahmias 2012: 12)

This does not mean that in Universe C people's mental states (their beliefs, desires, and decisions) have no effect on what they end up doing, and it does not mean that people are not part of the causal chains that lead to their actions. Rather, people's mental states are part of the causal chains that lead to their actions, though their mental states are always completely caused by earlier things in the causal chain that happened before them—if a person decides to do something in Universe C, then every time the universe is re-created with the same initial conditions and the same laws of nature, that person decides to do the same thing at that time and then does it. (Murray and Nahmias 2012: 18–19)

The subjects were asked whether a person in Universe C had “the ability to decide to do something other than they actually decided to do” or whether a “person has no choice about what they do.” In both instances, people gave a

compatibilist response: people in Universe C have the ability to do otherwise, and they have a choice about what to do (Murray and Nahmias 2012: 25).³²

These diverging results can be interpreted in different ways. However, as many people seem to believe that we can be responsible in the sort of universe as Universe C, which will be further discussed below, one interpretation is that even if the folk perception of agency is incompatibilist, the compatibilist conceptions of doing or acting might be what counts when discussing moral responsibility. Another interpretation is that people's reactions differ regarding what it takes to have a free will. The folk might judge free will to be compatible with determinism when confronted with deterministic conditions from a third-person perspective but not when confronted with determinism from a first-person perspective. If it is true that different judgments are made from different perspectives, we have to work out which perspective we should use, i.e., which perspective gives us intuitions with evidential value. The issue of free will is, however, not the topic of this dissertation. Therefore, I will now turn to the question about moral responsibility and the supposedly strong support by the folk of incompatibilism (regarding responsibility) and the skeptical arguments. I will first discuss the arguments from heteronomy and then the arguments from contrastive luck.

It has been widely assumed, by compatibilists and incompatibilists alike, that the incompatibilist position is the natural one. This is mainly due to the arguments of heteronomy. Non-philosophers have a strong intuition that if our actions are determined by factors outside our control we cannot be responsible for them. This is because we either need to be able to act otherwise in a very strong sense or because the sources of our actions ultimately need to be our own. Of course, things showed not to be that easy. Depending on how the cases are described, people's intuitions will differ (see Nahmias et al. 2007, Nichols and Knobe 2008, Feltz et al. 2009, Sarkissian et al. 2010, Murray and

³² For similar results, see for example Nahmias et al. 2007. For example, in one condition, 80 percent believed that people in that scenario had control, 82 percent believed they had free will, and so on (Nahmias et al. 2007: 227). See also Björnsson *forthcoming*. However, in a variant of the case, the result was the opposite: that people cannot decide to do otherwise (Murray and Nahmias 2012: 25). According to Murray and Nahmias, this is probably due to a misreading of the case, believing that people's motivational structures are "bypassed," which means that they are causally ineffective (Murray and Nahmias 2012: f.n. 27). (More on bypassing in section 3.2.3).

Nahmias 2012).³³ I will present the three most frequently discussed conditions that have been shown to generate different sorts of intuitions regarding excusing or non-excusing: high affect/low affect, concrete/abstract, and psychological/mechanistic, in order. Why I have chosen these three cases is simply because they are overwhelmingly the most frequently discussed when examining non-philosophers' intuitions regarding cases of heteronomy.³⁴

The most well-known study on concrete/abstract and high affect/low affect cases has been conducted by Shaun Nichols and Joshua Knobe (2007). In this study, subjects were first confronted with the following description of our universe:

Imagine a universe (Universe A) in which everything that happens is completely caused by whatever happened before it. This is true from the very beginning of the universe, so what happened in the beginning of the universe caused what happened next, and so on right up until the present. For example one day John decided to have French Fries at lunch. Like everything else, this decision was completely caused by what happened before it. So, if everything in this universe was exactly the same up until John made his decision, then it *had to happen* that John would decide to have French Fries. (Nichols and Knobe 2007: 669)

After this the subjects were confronted with either a concrete condition of high affect:

In Universe A, a man named Bill has become attracted to his secretary, and he decides that the only way to be with her is to kill his wife and 3 children. He knows that it is impossible to escape from his house in the event of a fire. Before he leaves on a business trip, he sets up a device in his basement that burns down the house and kills his family.

³³ Nahmias et al. 2005, Nahmias et al. 2006 found that non-philosophers were natural compatibilists. These findings were challenged, though, for instance by Knobe and Nichols 2007.

³⁴ It should be noted that a lot of things are happening in this area; new studies on the folk are being done, and new theories trying to explain the results are being put forward. Since my main motive with this theoretical introduction is to introduce my own research on the topic, published in 2013 and accepted the year before, I will try not to go into too much of what has been done in between when my paper was accepted and now in here. Instead, this will be presented in section 4.1.

Is Bill fully morally responsible for killing his wife and children?

Or a low affect and more abstract condition:

In Universe A, is it possible for a person to be fully morally responsible for their actions?

Nichols and Knobe found that 72 percent of the subjects judged Bill in the concrete condition to be responsible, whereas 86 percent stated that a person cannot be fully morally responsible in the abstract scenario (Nichols and Knobe 2007: 670).³⁵ This means that according to the vast majority of the subjects, people cannot be responsible in Universe A when the case is described in an abstract, low affect way. If Universe A is, however, described in a concrete, high affect way, a vast majority does think that people are responsible for their actions. These diverging results seem to hold on a cross-cultural level as well (Sarkissian et al. 2010: 352f). So the answer to the question whether the natural position is incompatibilist and whether the argument from heteronomy is a good one according to the folk, the answer is that we cannot know until we have some sort of theory with regard to which of these sets of intuitions, if any, we should give most credit.³⁶

Depending on how we describe the arguments from heteronomy as abstract/concrete or in a high affect/low affect way, we get different answers to our question whether people can be responsible when their actions are determined by factors outside their control. Something similar is true with regard

³⁵ For critique see Rose and Nichols *forthcoming*, Björnsson *forthcoming*. I discuss these in section 4.1.

³⁶ There is some reason to believe, however, that it is not affect that drives our intuitions here. Instead, it might be the level of concretion. First of all, in Murray and Nahmias (2012), the effect is also there when we simply compare low affect abstract and concrete cases; people judge those in the abstract scenario as less responsible than those in the concrete scenario (see for example Nahmias et al. 2006: 35–40 and Murray and Nahmias 2012). But this effect is less pronounced, which might mean that affect still does some work in our responsibility judgments. However, another study, by Cova and colleagues, found that patients suffering from “behavioral frontotemporal dementia,” meaning that their emotional reactions are impoverished, gave similar responses as ordinary people to the high affect and low affect cases (Cova et al. 2012: 860–862). This implies that people without high affect capabilities judge people in the high affect cases to be responsible while judging those in the low affect, abstract cases not to be responsible. This suggests that the driving force behind our intuitions in these cases might be the level of concretion and not the level of affect.

to psychological and mechanistic descriptions of cases. For example, Nahmias and colleagues (2007) constructed two deterministic scenarios where the only difference was that one was described from a neurological (mechanistic/reductive) and the other from a psychological (psychological/non-reductive) perspective in the following manner:

Most respected neuroscientists [psychologists] are convinced that eventually we will figure out exactly how all of our decisions and actions are entirely caused. For instance, they think that whenever we are trying to decide what to do, the decision we end up making is completely caused by the specific chemical reactions and neural processes [thoughts, desires, and plans] occurring in our brains [minds]. The neuroscientists [psychologists] are also convinced that these chemical reactions and neural processes [thoughts, desires, and plans] are completely caused by our current situation and the earlier events in our lives, and that these earlier events were also completely caused by even earlier events, eventually going all the way back to events that occurred before we were born.

So, if these neuroscientists [psychologists] are right, then once specific earlier events have occurred in a person's life, these events will definitely cause specific later events to occur. For instance, once specific chemical reactions and neural processes [thoughts, desires, and plans] occur in the person's brain [mind], they will definitely cause the person to make the specific decision he or she makes. (Nahmias et al. 2007: 224)

When asked whether people in the different scenarios were responsible, only 41 percent of the subjects judged people to be responsible in the neurological scenario, while 89 percent of the subjects judged people to be responsible in the psychological scenario. This result suggests the folk have incompatibilist intuitions when scenarios are described in a neurological (reductive/mechanistic) way and compatibilist intuitions when scenarios are described in a psychological (nonreductive/psychological) way.³⁷ Again, to the question whether the arguments from heteronomy hold, and whether incom-

³⁷ For a replication of this study with the same result, see Murray and Nahmias 2012. For studies on reductionism and responsibility ascriptions, see e.g., Green and Cohen 2004.

patibilism is the natural position, we have to have some sort of theory that tells us which set of intuitions we should trust and why.

Thus far, we have seen that the incompatibilist view, at least, is not straightforwardly supported with reference to common sense, as has commonly been thought. But as we saw in section 3.1.2, the main arguments today in favor of source incompatibilism are different sorts of manipulation arguments, such as Pereboom's four-case argument, which states that there is no principled difference between cases of heteronomy and of manipulation; since we judge people not to be responsible when manipulated, we should also excuse people whom we judge to be determined by factors outside their control. Accordingly, it might be interesting to make a minor digression to examine what the folk think about these sorts of cases. What Adam Feltz found when he tried the four-case argument on non-philosophers was that they believed that Plum was responsible in three of the four cases (Feltz 2013: 52–59). Only in the case of direct manipulation did the folk judge Plum not to be responsible. However, it should be noted that the responsibility ascriptions were provided on a falling scale, where direct manipulation was to the largest extent seen as an excuse, followed by indirect manipulation, culture, and finally determinism. So in the case of determinism, for instance, Plum got about 5.7 on a 1 to 7 scale, and in the case of intentional direct manipulation, Plum got about 3.8 on the same scale.

Feltz argues that these results show that Pereboom is wrong: the folk *appreciate* the difference between the direct manipulation cases and that of determinism, which means that Pereboom cannot use laypeople's intuitions to support his claim (Feltz 2013: 59f).³⁸ But as far as I can see, Pereboom's suggestion is not that the folk will not react differently to scenarios of deterministic universes and to scenarios where people are manipulated by other agents. Instead, what he says is that ordinary people have a strong intuition that people who are manipulated are not responsible and that there is no relevant difference between those cases and those of determinism. Hence, those people *should* find determinism to be a problem for responsibility. That is, those who have the intuition that we can be responsible in a deterministic universe have

³⁸ For another study on manipulation with purposely compatibilist results, see Sripada 2012. For critique, see Björnsson unpublished manuscript.

to explain why we are responsible in the deterministic universe when we are not in the manipulation case. However, to have a unified account of folk intuitions with regard to responsibility, we need to explain the results in the Feltz study, such as the falling degree of responsibility ascriptions when the intentionality of the manipulation increases.³⁹

A lot of interesting work has been done, then, on the case of heteronomy. Our responsibility judgments depend on a range of different and seemingly irrelevant factors such as whether they are described in a high affect or low affect way and so on. The other sort of skeptical argument put forward in the responsibility debate is that of moral or contrastive luck (see 3.1.3). There has not been done nearly as much research on the problem of moral luck, but there is a growing interest in this subject as well. For example, Young and colleagues have examined people's intuitions regarding resulting moral luck (Young et al. 2010). The case they used was Mitch, who on his day off is going to give his two-year-old a bath. When the bath fills, the phone rings, and Mitch goes out to take the call, leaving his son behind in the bathroom without emptying the bathtub. Without going into the details of the case, when he comes back, he either finds his son "face down" in the tub, in the tub bathing, or beside the tub as he was when Mitch left. In these scenarios, people judged Mitch to be more blameworthy in the case where his son was "face down" than when he was still beside the tub (Young et al. 2010).⁴⁰

What seems to be the case, then, is that people share the intuition by Nagel and many other philosophers above: people are more blameworthy when things happen rather than not happen. But, Nagel would say, the more blameworthy individual, in our case Mitch whose child drowned, has suffered from bad (resultant) luck. There is no difference between his thoughts, intentions, beliefs, etc., and those of Mitch whose child waited at the tub. So, according to Nagel, it seems that on the one hand, we have the intuition that "bad luck Mitch" is more blameworthy than "good luck Mitch," but on the

³⁹ What receives support from this study, however, is Mele's hypothesis that the explanation for our exculpating intuitions in the Plum cases is that of manipulation and not deterministic causation (Mele 2006: 138–144).

⁴⁰ The version of the paper I have read is open source and with no pagination, which is why I have left the page number out. However, the description of the case is on page 5, and a graph describing the blameworthiness can be found on page 9.

other hand, when we compare the two, we get an exculpating intuition: bad luck Mitch is perhaps not that culpable after all. Whether the folk share this latter exculpating response, however, is not examined in this study. The same is true with the studies I am going to discuss in the next section, which is a weakness and should be amended.

Another finding much discussed among experimental philosophers is the “Knobe effect” (Knobe 2003).⁴¹ This effect is similar to that of circumstantial moral luck.⁴² In a well-known illustration, an agent has similar dispositions in two cases—he or she wants to make money and does not care about anything else, but is only judged to be responsible in the case where his or her actions have negative side-effects.⁴³ The case that Knobe used in his original experiment and which is still used when doing these experiments is the following:

The vice-president of a company went to the chairman of the board and said, “We are thinking of starting a new program. It will help us increase profits, but it will also harm [help] the environment.”

The chairman of the board answered, “I don’t care at all about harming [helping] the environment. I just want to make as much profit as I can. Let’s start the new program.”

They started the new program. Sure enough, the environment was harmed [helped]. (Knobe 2003: 191f)

For example, Cole, Wright, and Bengson (2009) found that when the environment was harmed (“harm condition”), 92.6 percent of the participants judged that harming the environment was bad, and 88 percent believed that the chairman deserved blame for doing it (Cole, Wright and Bengson 2009: 30). When instead the environment was helped (“help condition”), 90 percent judged that helping the environment was good, but only 14 percent thought

⁴¹ For an overview of the discussion about the Knobe-asymmetry, see for example Petit and Knobe 2009.

⁴² It should be noted that neither Joshua Knobe nor anyone else, as far as I can tell, is claiming that these cases are cases of moral luck. But regardless of that, there are relevant similarities that I will discuss when describing the Knobe-asymmetry below.

⁴³Knobe was, and still is, interested in the intentionality of side effects and not primarily in responsibility. That people assign responsibility in the same way that they assign intentionality and praise/blame is something we have found in a study of our own (Björnsson and Persson 2012: 14; f.n. 25).

the chairman deserved praise (Cole, Wright and Bengson 2009: 31). Presumably, the reason why people differed in their judgments of praise and blame has to do with their responsibility judgments. The chairman was judged to be responsible in the harm but not in the help condition. Of course, this asymmetrical result is not something the conceived theories of responsibility can account for; according to those, the chairman is equally responsible in both cases.⁴⁴

3.2.3 Explanations of diverging results and why they fail

There is a range of accounts trying to explain all these diverging results in the responsibility area. First of all, there are those who argue that there is no uni-

⁴⁴ Another finding in the area of experimental philosophy and responsibility is that the folk might even think that people are sometimes responsible for things over which they have no control (Woolfolk et al. 2006). According to these results, what might be important instead is the idea of identification. The case that Woolfolk et al. (2006) use in support of their conclusion is one with Bill and Frank, who have been captured by terrorists. In all of the versions of the case, Bill kills Frank, with different degrees of identification and constraint. According to the most extreme case, and for us the most interesting, Bill is subjected to an absolute constraint while at the same time, he highly identifies with the action performed. The absolute constraint is the effect of a “compliance drug” that the terrorists give him. This drug makes it impossible for his body to resist any orders given. When his body moves on orders, he does not feel it move, and he does not try to move; he feels like a puppet. At the same time, in this version of the case, he wants to kill Frank because Frank has had an affair with his wife. Accordingly, when the terrorists order him to kill Frank, although he does not feel his body move, he feels no reluctance that this should happen, because this was something he wanted to happen.

Here, people believe that Bill is responsible to a quite significant extent, 3.25 on a 7-point scale (Woolfolk et al. 2006: 297). This might imply several different things. For one, we do not have to be able to do or act otherwise to be responsible, according to the folk. It is enough that we identify with our actions. For another, it does not seem relevant how our actions have been produced. After all, Bill’s action is caused by the terrorists and by their actions, not by Bill himself. Together, these two implications rule out many of the most renowned theories on responsibility that we have discussed thus far. This is true about both source views as well as leeway views. The case with Bill and Frank has these implications for the theories on responsibility because we do not have to be able to act otherwise, and the sources of our actions do not have to be “unmanipulated,” etcetera.

However, there are many problems with this study. First, it has been difficult to replicate the results when the case is described somewhat differently. For example, when Alfred Mele conducted a similar study, only 21 percent judged the person being under absolute constraint to have free will (Mele 2012: 430). Since free will among many is believed to be necessary for responsibility, this casts the results of Woolfolk and colleagues in a bit of a different light. Second, the interpretations of the results are a bit inflated. It was found that people actually do not believe that the compliance drug worked in an absolute fashion; they thought that Bill had some leeway in what he did. Another finding was that people in a similar way hold others responsible for events over which they have no control when they disapprove of some decision they made earlier. All of this suggests that even though Woolfolk and colleagues at first glance seemed to have found something interesting, it does not hold under scrutiny.

fied folk theory of responsibility; instead, the theory is contextualist (Woolfolk et al. 2006: 298f) in the sense that the theories differ from situation to situation and culture to culture. Others have argued that different people have different concepts of responsibility (Knobe and Doris 2010, Feltz et al. 2009, Feltz 2013). Now, even though this might be true, the conclusion seems premature. After all, the experimental research on the subject of responsibility is still quite new, and it is reasonable to think that we might find a unity in these, on the face of it, disparate intuitions.

Another explanation for why we judge people to be responsible under certain deterministic conditions has to do with our emotional responses. According to the *affective performance error model* devised by Nichols and Knobe (Nichols and Knobe 2007: 672), strong emotional reactions to certain cases make us commit errors when thinking about questions on responsibility.⁴⁵ For example, when we read about Bill killing his family to be with his secretary, we react to this very strongly, so that we forget the matter of Bill's being determined, and because of this, we judge him to be responsible. If we had thought about it in a cool hour, however, we would judge him not to be responsible.

There might be something to this argument. However, it cannot give us the whole picture. First, there are many cases where laypeople judge people whose actions are determined by factors outside their control to be responsible, even though these cases are low affect (see for example Nahmias et al. 2006, 2007, Murray and Nahmias 2012, Cova et al. 2012). People judge agents whose actions are determined by factors outside their control to be responsible for good things, such as contributing to charity, as well as for minor moral transgressions, such as stealing jewelry in a jewelry store. Second, the affective performance model cannot explain the difference between concrete and abstract cases. Many of the concrete cases are either morally neutral or morally positive, so affect cannot account for the difference in responsibility judgments. Last, it seems that even though we would come to accept that

⁴⁵ For additional support, see Roskies and Nichols 2008. They find that people are more likely to ascribe responsibility to those who are portrayed as living in our world rather than in another deterministic world. The conclusion is that we are biased due to emotional responses when being confronted with our world and are hence less trustworthy when it comes to judgments about these cases. However, see footnote 36 for a rebuttal of this view.

affect makes a difference in responsibility judgments, we would want to know why.

A third way of trying to explain the divergent results with regard to cases of heteronomy is that we conflate bypassing with heteronomy and determinism (Nahmias et al. 2007, Feltz et al. 2009, Murray and Nahmias 2012). This is called the “Error-Theory” (Nahmias and Murray 2010), but since there are other error-theories out there, we can call this the “*Bypass Hypothesis*.” When an individual’s motivational structure is bypassed, her motivational structure lacks causal efficacy. This can be true for a number of different reasons. For example, if fatalism is true, then whatever we do, the things that are “fatalistically determined” will happen no matter what. If I am fatalistically determined to die on November 16, 2022, then it does not matter whether I exercise and eat well or try to be on the lookout for dangers on that day. I will still die on November 16, 2022. Or to take a more mundane case, if an addict is fatalistically determined to continue to be an addict, then she will continue to be an addict no matter what. It does not matter whether she tries therapy, medicine, or something else. She will still be caught in her addiction. Now, fatalism is different from heteronomy and determinism. For example, whether your actions are determined by factors outside your control has nothing to do with whether you can lose your addiction. Of course, it is not ultimately under your control to lose your addiction. That depends on factors outside your control. But we might assume that if you try hard enough, giving therapy and medications a real chance, you will become a former addict. Nothing in determinism per se precludes that from being true, which means that determinism and bypassing come apart.

That laypeople have incompatibilist intuitions because they conflate heteronomous agency and that of being bypassed has, as we saw above, been questioned (see footnote 35). However, even if it turns out to be false that bypass judgment can explain exculpating judgments, there are other deficiencies regarding the bypass hypothesis. First, it cannot explain all the things that the proponents of the view want it to explain. The example the proponents of this hypothesis cite is the diverging responses with regard to some of the abstract and concrete cases (Murray and Nahmias 2012: 21). After controlling for affect, people in the concrete condition still judge those whose actions are

determined by factors outside their control to be more responsible than when people are in the abstract condition. Second, the bypass hypothesis cannot account for the Knobe effect. It should be noted that the proponents of this model have no ambitions of solving the Knobe-asymmetry with their model, but it would still be preferable to have a model that could account for all, or at least most, of the perplexities regarding responsibility judgments that have arisen as a result of the findings in experimental philosophy.

Last, when discussing the Knobe asymmetry, philosophers have tried to explain it in different ways by referring to our normative expectations (Nelkin 2007, Petit and Knobe 2009, Knobe *forthcoming*). For example, there is a difference in moral status in doing and allowing. Some people say that I am allowed to let someone die (allowing) but am not allowed to actively kill the same person (doing). I do not have to save the drowning child, i.e., I can allow the drowning to happen, but I am not allowed to drown him by holding him under water until he stops breathing. Translating this view into the case with the vice-president, he is allowed, people might think, to let the environment not be helped, but he is not allowed to actually harm the environment, even though this is just a side effect of what he is doing. That is, it could be argued that in the harm scenario, he does something morally bad and hence does not live up to our moral expectations. However, in the help scenario, he does not do anything morally good (or bad) and hence does not violate any of our normative expectations. This suggests that if this model is correct, we should judge the vice-president to be responsible in the harm but not in the help scenario, which, as we know, is the case.

I will not discuss this idea at any length here. Needless to say, the idea is too narrow to explain the other phenomena discussed in this section, and an explanation for how and why these expectations play this role is also still to come. So even though there seems to be some truth to this hypothesis, it needs to be more thoroughly worked out, and it has to be widened so that more things can be explained.

3.2.4 Summary and outlook

To sum up: different heteronomy and moral luck arguments have been tested, and the results are perplexingly divergent. Non-philosophers react in a wide

array of different and unexpected ways, depending on how these cases are described, and no unified theory explaining these judgments is in sight. More precisely, what we would need is a theory that can say why the folk judge people to be responsible in the:

- (1) ...concrete but not the abstract condition.
- (2) ...high affect but not the low affect condition.
- (3) ...psychological but not the mechanistic condition.
- (4) ...harm but not the help condition.
- (5) ...ordinary deterministic but not the direct manipulation condition.

Some progress in answering these questions has been made through the affecting performance model, the bypass hypothesis, and the idea of normative expectations. However, as we have seen, all of these views have decisive deficiencies, and they are by no means anything near a unified theory about responsibility judgments. All in all, then, much interesting work has been done and results found, but there is still much more work to do before we can be satisfied.

PAPER 2: A UNIFIED EMPIRICAL ACCOUNT OF RESPONSIBILITY JUDGMENTS

In Paper 1, we showed that the explanation hypothesis can account for why we come to different judgments when confronted with different arguments. For example, when we are confronted with an abstractly portrayed regress argument, we tend to take a certain explanatory perspective where an agent's motivational structures do not stand out as significant explanations, which accounts for why we judge them not to be responsible. The motivational structures are just links in the causal chain. In this paper, we answer question 2, showing that the explanation hypothesis can resolve a number of perplexing phenomena discovered by experimental studies of responsibility judgments, mainly items 1–4 listed in the summary 3.2.4, but also that the explanation hypothesis can stand being tested directly. (Item 5 will be discussed in 4.2.1.)

We begin with one study where the explanation hypothesis was tested directly (Johansson 2010). In this study, subjects read a number of different scenarios and were asked if they thought the people in those scenarios were responsible for certain outcomes and if their motivational structures could explain the outcome portrayed in the case. In one scenario, Kevin sees a dog in a car on a hot day. He understands that it will suffer from the heat but decides that it is none of his business. After a while, the dog passes out due to the heat. Subjects are asked both whether Kevin was responsible for the dog passing out and whether it happened because Kevin did not care enough. Based on the explanation hypothesis, one would expect answers to the two sorts of questions to be quite strongly correlated. This was indeed the case.

The explanation hypothesis can also give a unified account of questions 1–5 posed in the summary 3.1.4, above. There are similar answers to why we react as we do when confronted with the different arguments from heteronomy, which where: abstract vs. concrete (1), high affect vs. low affect (2), mechanistic vs. psychological (3).

Many of us believe that actions, decisions, etc., are straightforwardly explained by people's motivational structures and that these structures cannot be straightforwardly explained by something else. So when asked about a deterministic scenario where the agent's motivational structure is a salient part of the description, it is not strange that people judge that the agent is responsi-

ble. In abstractly (mechanistically) described cases, we are invited to take an abstract and non-everyday explanatory perspective where agents' motivational structures fall into the explanatory background. This is because the motivational structure is not described at all (or at least, not in an ordinary way) in combination with the fact that it is explicitly said that all events are "fully caused" by "what happened before," which leads us to expect the agents' motivational structures to be straightforwardly explained by something else; hence, they cannot significantly explain anything.

Conversely, when confronted with concretely (psychologically) described cases with explicit mention of actions and motivational structures, the case is radically different. We usually expect motivational structures to straightforwardly explain actions without anything straightforwardly explaining the motivational structures. This suggests that when we have a motivational structure to lock on to, when trying to find an explanation for the event at hand, this leads us to take an everyday explanatory perspective, judging the individuals under the concrete condition to be responsible even though their actions are determined by factors outside their control. In the high affect/low affect cases, this becomes even more prevalent. In the high affect cases, people make grave moral transgressions. This is something we expect them not to do; hence we are imbued with an everyday perspective that implies that their motivational structures become a significant explanation. In the low affect cases, the reverse is true. Because agents do nothing out of the ordinary, we take an abstract explanatory perspective; hence their motivational structures do not stand out when searching for a significant explanation.

The *Explanation Hypothesis* can also give an answer to item 4 in 3.2.4: why people give asymmetric responsibility ascriptions for known side effects (which I discussed under the heading "moral luck"). First, it seems natural and straightforward to say that the chairperson did the environment harm because he or she did not care enough about it. This is because indifference toward valuable things is normally believed to harm them. It seems much less natural and straightforward to say that the environment was helped because the chairperson wanted to make a profit. Second, we have normative expectations of people not to harm the environment. This implies that the "not caring" part about the environment stands out as a significant explanation: we believe

not caring about the environment harms and that this is morally wrong. Thus the explanation hypothesis accounts for why we judge the chair to be responsible in the harm condition. Compare this to the benefit condition, where the chair did not oppose performing actions benefiting the environment. Not opposing helping the environment is something we do (normatively) expect of people, implying that the chair's motivational structure falls into the background, which, according to the explanation hypothesis, accounts for why we judge the chair not to be responsible for benefiting the environment.

All in all, if what we argue in this paper is correct, we have strong support for the explanation hypothesis, which in turn lends credibility to the notion that the question whether the arguments from heteronomy and luck ultimately give us reason to believe that incompatibilism is true is a normative or metaethical question.

3.3 RESPONSIBILITY, ACCOUNTABILITY, AND ILL HEALTH

Thus far we have discussed under which conditions individuals should be held responsible for their ill health by examining the philosophical debate and people's intuitions. Here we will discuss some normative considerations relevant to the issue. It should be noted that this section is distinct from 3.1 and 3.2 because its aim is different. In 3.1 and 3.2, the purpose was to clarify the reasons for asking questions such as (1) why we receive skeptical intuition when confronting skeptical arguments and (2) why we need a unified explanation for people's intuitions. Posing these questions led us to discover that the current discussions had reached different forms of stalemates (3.1) or that the current theories explaining people's intuitions were deficient or too narrow (3.2).

The intention here is not to show that different normative theories are false or that discussions about these theories have reached impasses. Instead, to advance to the overarching question about whether we should hold people responsible for their health, I choose the narrow route (see section 2 above), applying the most common normative considerations from relevant areas (mostly bioethics and healthcare ethics) to the questions at hand. What I propose, then, is to comment on the different considerations and normative ethics in general, so that we can understand the context before dealing with questions 3 and 4 in the summaries below.

3.3.1 Some introductory clarifications

To begin, many different views of when we should attribute responsibility for some outcome or act to people, or when they should be held responsible for an act or an outcome, exist in normative ethics, with a vast range of intricacies. These accounts are forward-looking and/or backward-looking, substantial and/or procedural, monistic and/or pluralistic, and fundamental or derivative, which will be discussed in order. I will finish off by saying something about what "holding people responsible" might amount to.

When deliberating on which institutional practices we should endorse, there are backward- and forward-looking reasons. Forward-looking reasons are typically related to consequentialist theories. For example, we might think

that well-being is the only issue that matters, and what makes an act right depends solely on whether it creates more well-being than any alternative. If this assumption is true, we will only have access to forward-looking reasons when justifying our institutional practices; this means that the relevant fact about what makes a certain institutional framework better than another relates to the effects of this framework.⁴⁶ Of course, we might think that this view is wrong or that it does not give us the whole picture. Instead, we might believe that other factors are important from the viewpoint of justice. It might be argued that for our institutional practices to be just, they have to be sensitive to people's choices and what they deserve. For example, if I performed an evil act in the past, then I deserve some sort of negative response in the present or in the future. However, we can only know what someone deserves by looking backward, which is why these sorts of justifications are essentially backward-looking.⁴⁷ Many people believe that backward-looking reasons are essential when creating just, health-related institutions. I will discuss these views under the umbrella, "responsibilitarianism," below.⁴⁸

⁴⁶Not everyone would agree; see e.g., Tännsjö 1998 for a different view.

⁴⁷We have to be careful here, though, when talking about desert- and choice- (desert for short) sensitive institutions. There are both institutional and pre-institutional deserts. To give individuals what they "institutionally deserve" simply refers to what they are entitled to receive according to the social arrangements at hand. For example, if we have public-funded healthcare where everyone has a right to receive help when in need, then people in need institutionally deserve to get help. (The discussion about institutional desert is more refined than this. For references, see e.g., Olsaretti 2003). These institutional frameworks could of course be defended on forward-looking grounds, which is why philosophers skeptical to desert, such as John Rawls (1971), maintain that people should get what they deserve. The idea is just that what they deserve and how they come to deserve it are constituted by institutional practices justified with regard to other reasons.

To give people what they "pre-institutionally deserve" is instead to give them something based on whether they have done something good (moral desert), prudent (prudential desert), etc. That is to say, we should give the deserving what they deserve because they deserve it, and this should be the basis for justifying our institutional practices. Thus, we give a backward-looking justification of why murderers deserve punishment, not because we presuppose an institutional framework where murder is illegal and punishable, but because murder is an evil act that deserves punishment; hence, it is the reason why we should have a law against it.

⁴⁸When discussing substantial responsibility, then, we want a notion of responsibility substantial enough to make it permissible to hold people responsible on backward-looking grounds. Many philosophers believing in free will and moral responsibility also deem that holding us responsible on the grounds of desert is unfair (see e.g., Persson 2005: 409–429). It is permissible to hold us responsible on forward-looking but not on backward-looking grounds, because we cannot be responsible in the substantial sense needed. We can only be responsible in looser senses, for example, in an attributive sense, which roughly

Substantive and procedural accounts are also relevant to the question of whether people should be held responsible for their ill health. Substantive accounts relate to the pattern of the good. For example, according to the sufficiency view (e.g., Frankfurt 1987), if people cross a certain threshold securing goods of special importance, the outcome is just. Alternatively, based on the utilitarian view, a just outcome constitutes the maximal amount of well-being in comparison to other alternatives (e.g. Sidgwick 1907, Tännsjö 1998). Of course, there might be different procedures for us to guarantee that the outcomes contain the maximal utility or that everyone is above the threshold. But these procedures are just instrumental in the sense that if they did not guarantee, for instance, the maximal amount of well-being, then they would not produce just results according to the classic utilitarian criterion of rightness.

Procedural accounts state that whether an outcome is just links to how the outcome is created, not the outcome itself. The most well-known procedural account in the context of priority setting is *Accountability for Reasonableness* (A4R) by Norman Daniels and James Sabin (2008).⁴⁹ According to Daniels and Sabin, the problem is that a wide range of plausible, competing normative considerations allows no consensus on the correct one, even among reasonable people. It might be responsibilitarianism, utilitarianism, egalitarianism, none of these, or a combination of them. Daniels and Sabin suggest that the solution is to create a fair and legitimate decision-making process most people can support, irrespective of their favored consideration. The authors argue that if the process is fair and legitimate, the outcome will be such as well. The conditions for a fair process are that the rationales for the decisions are made public, the rationales are reasonable, the decisions can be challenged, and some sort of regulation guarantees that these conditions are met (Daniel and Sabin 2008: 45). Although this approach has been widely popular in recent years, I will not use it when discussing questions 3 and 4, whether policymakers and healthcare personnel should perceive people as responsible for their ill health and whether we should prioritize bariatric surgery, respectively. The reason is that even if these approaches are sound, we still need substantial arguments

means that our actions reflect our motivational structure (Watson 1996). However, this does not imply that we are also responsible in a substantial sense, which justifies reproach for what we have done.

⁴⁹For other procedural accounts, see e.g., Fleck 2009.

building on normative and empirical facts when discussing matters of holding people accountable in general and how to prioritize our resources in particular, e.g., within the framework of a procedural solution of the mentioned type.⁵⁰ I will focus on producing such arguments here.

There are also monistic and pluralistic accounts. According to the monists, one master value trumps all others; in contrast, the pluralists argue for many values that have to be weighed against each other. Most people engaged in the debate are pluralists, and it is easy to see why. For example, responsibilitarianism has been accused of being too harsh (“the harshness objection”).⁵¹ When an individual is responsible for negligence leading to his or her severe injury, responsibilitarianism implies no obligation on our part to help the injured individual, even if he or she is dying without our help and we could give it without any sacrifice on our own. Many people believe this view to be absurd, which counts as a strong argument against it. However, this is easily solved if we grant that there are other values than giving people what they deserve. When an individual’s life is at stake, and we can help without sacrificing something that comes close to the value being lost by not helping, we have an obligation to help because of the enormous value of a human life. This position holds true even though, in many other cases, we do not have to help when the person is responsible for his or her troubles.

When discussing whether we should hold people responsible for their health, the different normative considerations being used are also fundamental or derivative. Fundamental normative considerations (or theories) are ideas such as utilitarianism and egalitarianism. It is not that we hold utilitarianism to be true because we believe some other normative consideration to be correct. Instead, it is fundamental in the sense that all other considerations we use in day-to-day life, for instance, rules of thumb, decision-making models, and institutional frameworks, ultimately have to be supported by the fundamental consideration. The fundamental normative considerations in the discussion about people’s accountability for their ill health and priority setting

⁵⁰Together with Bengt Brülde, I have argued for using A4R when prioritizing in public health work (Brülde and Persson 2011).

⁵¹This is also called “the negligent victim objection.” For a discussion, see e.g., Anderson 1999; for further references, see Duus-Otterström 2012. Some people have also tried to answer this objection; see e.g., Knight 2009, Segall 2010, Duus-Otterström 2012.

are responsibilitarianism, utilitarianism, different types of egalitarianism (sufficiency, priority, equality), and various autonomy views. From these perspectives, other more precise rules of thumb and so on have been derived. Those often used by healthcare personnel on the “shop floor” are beneficence and maleficence, but also that from greatest need.⁵² I will discuss all of these considerations (fundamental and derivative) in order, but before that I will say something on what holding people responsible might amount to.

Last, before we go on discussing the normative theories, holding people responsible (accountable) is something we could do *ex post facto* and *ante facto*.⁵³ To hold people responsible *ex post facto* is to hold them responsible *before* the outcome has occurred, which often means holding them responsible for their risky behavior. There are many ways we could do this. One is to make people pay for what their projected healthcare costs will be, say, if they indulge themselves in a certain behavior. Putting taxes on cigarettes and forcing climbers and caretakers of dangerous animals to take extra insurance are examples of this. Another way to hold people responsible post fact is to give them lower priority in priority contexts. For example, people who continue drinking while being on the live transplant list could be forced to wait longer.

To hold people responsible *ante facto* is to hold them responsible *after* the event has occurred, which might mean holding them responsible for actions/omissions and/or the event of injury or ill health. This can be accomplished in a variety of ways as well. One way is to make them pay for all, or some, of their healthcare costs after they have obtained treatment in cases when society (or the insurance companies) normally pays. Another way is to give people lower priority from the beginning. That is, people who have developed lung cancer as a result of smoking should be given a low priority from the start, even if they stop smoking. A third way of holding people responsible for their ill health is for healthcare personnel to blame them. This is probably quite common. Most people have probably had a doctor reproach them for taking unnecessary risks, for having high blood pressure, etc.

These ways of holding people responsible can be justified on different grounds, using the fundamental or derivative principles explained below. For

⁵² The principle of greatest need is also used at higher levels by policymakers when setting priorities.

⁵³ For different ways in which we can hold people responsible for their ill health, see Martin 2001.

example, we might have utilitarian reasons for blaming people who unnecessarily put their health at great risk. This is because subjecting them to reactive attitudes, we might assume, will make at least some of them quit, which in turn will create a better outcome than the alternative of not blaming them. However, using the same line of (utilitarian) reasoning but assuming that people do not react on incentives in this way, we may have reason for not holding them responsible even if they are responsible. The underlying principle(s) will also determine to what extent we should hold people responsible when holding them responsible is justified. Some people seem to think, for instance, that when people are responsible for their ill health, they should pay for all of its costs. But of course, this may not be true. Again, using utilitarianism as the example, we might assume that having people pay for some of the costs when they are responsible has good effects, because it gives people incentive to take better care of their health. But if we push this too far, it might be the case that the effects on balance become negative. Those who get sick or injured, as many uninsured do in the United States, might propel downward with deteriorating health and no ability to pay for the increasing costs of their medical bills. This in turn will probably have disastrous consequences for those afflicted, with no matching positive consequences overall. So from a utilitarian standpoint, assuming that these speculations are true, it might be a good thing to find a position in midway, holding people somewhat responsible even when they are fully responsible.

When I discuss whether we should hold people responsible for being obese, the “holding responsible” part is meant to be through giving them low priority in priority-setting contexts. This and the other ways of holding people responsible, such as having people pay for their healthcare by themselves or not setting any preventive work measures, is something I use when I discuss what it is reasonable to expect for policymakers and healthcare personnel to do if they believe people are responsible.

3.3.2 *Responsibilitarianism*⁵⁴

The fact that we are allowed to hold people responsible for their ill health on the grounds that they are responsible for it is intuitive to lay people and many philosophers (Knowles 1977, Rakowski 1991, Knight 2009, Segal 2009, Kagan 2012). For instance, it seems unfair to force me to pay the healthcare bill for someone who has not bothered eating and exercising properly. I should not be punished for their misconduct. Similarly, it seems problematic that individuals who take calculated risks, living “the good life” and ingesting drugs or excessive amounts of alcohol, should receive the same priority (e.g., for a liver transplant) as myself, who has been deliberately moderate with these sorts of activities. I am in this position just because of bad luck; they are here because they thought it was worth it.

Responsibilitarians all agree that we have reason to hold people accountable when they are (substantially) responsible for different outcomes. However, there are different views on what further grounds in general, and regarding health in particular, we should do so.⁵⁵ Three different grounds have been proposed: choice, desert (Arneson 2011: 34–36), and restoration (Smart 1994, Walker 2010). According to the choice view, we have reason to hold people responsible because we own, so to speak, the consequences of our actions, i.e., we do not hold people responsible because they deserve it, because they have done something good or bad. Instead, it is similar to gambling; sometimes people win, sometimes they lose, but in both cases they are entitled to their winnings or losses.⁵⁶ In contrast, the desert view posits that the reason for responsibility is connected with something (morally, prudentially, etc.) good or bad. For example, it might be morally wrong to take great risks with one’s health because, we might assume, it might cause harm to others by unnecessarily using up the scarce resources available. When this is true, we have reason for holding an individual responsible by, for instance, giving the liver to someone else. Finally, according to the restoration view, people who have been negligent with their health and thereby have created scarcity in

⁵⁴ It should be noted that Carl Knight (2009) uses the same term, but he means something different than I.

⁵⁵ This might suggest that responsibilitarianism, after all, is not fundamental but derivative, which might be so, but I will not discuss it here.

⁵⁶ See for example Ronald Dworkin (1981) on option- and brute luck.

healthcare resources should be held accountable through the same measures already described. Presumably, the restoration view is not a fundamental consideration but should instead be understood as being derived from the choice or desert view.

It is important to be clear about what the grounds are for holding people responsible, because it answers some important objections against responsibilitarianism, and it might also have some policy implications. One argument against responsibilitarianism can be called “the argument from arbitrariness” (e.g., Wikler 2004, Segall 2009).⁵⁷ This is a dilemma which many people think is impossible for responsibilitarians to solve. The first horn of the dilemma is that holding everyone responsible for their self-inflicted injuries or ill health leads to absurd implications. Having firefighters pay for their healthcare needs when they have injured themselves in the line of duty is just one example. The second horn of the dilemma is that if we do not hold everyone liable for their injuries or ill health when they have some responsibility for the outcome, then we act arbitrarily. We have no principled reason to hold one but not the other accountable, provided that they both fulfill the relevant responsibilitarian condition. Instead, what steers our judgments here constitutes prejudices against poor people who may not have the same refined tastes as those of the middle class and who, because of it, eat much more fatty and sugary foods. To this, a desert-based responsibilitarian, for instance, might reply that firefighters should not be held responsible because they do not deserve it. They have done well and deserve, if anything, to be rewarded. This position means that we have a principled way of excluding those who intuitively should not be held responsible, while holding others liable.

Policy implications of the different views might also be quite significant. For example, take the difference between a choice-based and a desert-based view. According to the choice-based view, an individual owns the consequences of one’s actions, which might mean that one has to pay for healthcare needed when one is responsible for injuring oneself. It involves no discussion about punishment and proportionality. When deciding what one owes, if anything,

⁵⁷Neither Wikler nor anyone else, as far as I know, has formulated this argument in the way I present it here. My description condenses and combines several common arguments often used in the debate about holding people responsible for their ill health.

it is just a matter of examining whether one is responsible for the event at hand. On the other hand, according to the desert view, we may have to consider the matter of punishment and proportionality. We hold a person responsible because he or she has put someone else at risk, but since we do so on the basis of desert, it is important not to give the person more than he or she deserves.⁵⁸ It would be unfair.⁵⁹ This leads to the institutions not only having to look into the degree of responsibility of the individual but also to what extent she should be held responsible so that the proportionality requirements are met. The point is, depending on what view is correct, choice or desert, different policies might be implied with which we have to be careful when defending responsibility-sensitive institutions on the basis of responsibiliarianism.⁶⁰ Of course, if people are not responsible, then this will not be an issue.

A wide range of discussants believe that, generally, people are not responsible for their ill health for reasons other than those discussed in section 3.1; hence, they maintain that we should not hold people responsible on such grounds (see e.g., Minkler 1999, Wikler 2004, Cappelen and Norrheim 2005, Daniels 2008). Two of these reasons are presented here, which seem to be the most prevalent ones when discussing obesity. The first relates to the control condition, and the second is associated with the epistemic condition (Aulisio 1996, Benjamin 1997, Olsen 1997). We will see that they are not obviously excusing conditions but something we have to examine to find out if they apply to the scenarios involved when discussing responsibility and health.

To be responsible for anything, we should have some control over it. How this notion of “control” should be unpacked is an even more complex ques-

⁵⁸For an interesting discussion about the geometry of desert, see Kagan 2012.

⁵⁹For an argument of this sort, see Harris 1995 and Wikler 2004. It is often combined (which Harris and Wikler also do) with the view that it will be extremely difficult to discern who is responsible for what and to what extent, which means that we will punish many innocent individuals too harshly. However, I am not convinced by this argument. First, most normative theories share a similar problem, to allocate resources to those who should get them, because it is difficult to know who these people are. For example, if one should allocate resources to those with the greatest need, then one has to pinpoint who they are, which is not easy. Second, trying to give people what they deserve might increase the chances of doing so to a larger extent. If giving people what they deserve is a good thing, we should aim at giving them more rather than less of it; hence, it might still be viable to build responsibility-sensitive institutions.

⁶⁰Theoretical issues might also be addressed in different ways, depending on the view. For example, the desert view might have an easier time handling the argument from arbitrariness, but a more difficult time dealing with our intuition that people should not be punished for becoming ill.

tion than the discussion in section 3.1 suggested. The concept might also be connected with external and internal resources, such as having enough money, having access to the right information, being rational, and so on (e.g., Mele 1995, Fischer and Ravizza 1998, Dennett 2003). Here the focus is on a widely held objection against people typically being responsible for their ill health in general and obesity in particular, which concerns addiction (Higgs 1993, Olsen 1997, Minkler 1999, Tonti-Filippini 2004, Wikler 2004, Daniels 2008).⁶¹ This case can be true for many reasons such as self-deception (Levy 2003, Walker 2010b), irrationality (Waal and Mörland 1999), fear of the powerful experience of withdrawal (Watson 1999), or lack of impulse control (for references, see Levy 2006, Foddy and Savulescu 2010). The most common excuse is the one from impaired impulse control, which is why I am going to focus on this criterion.

It seems then that addiction might be an excuse, because it generates a lack of impulse control. There are many ideas about how this criterion should be spelled out, whether it has to do with not being able to act in accordance with our higher order wills (Frankfurt 1971, 1988), not being moderately reason-responsive (Fischer and Ravizza 1998, see also footnote 61), not having an extended agency (Levy 2006), and so on. However, I will not go into how the impulse control criterion should be spelled out in any greater detail. Instead, I

⁶¹ However, others think that addiction is no excuse (Wallace 1999, Watson 2004). A commonplace position in the philosophical debate on responsibility is that as long as we are moderately reason-responsive, i.e., that we would have made another choice in a different situation with the same motivational structure, we are responsible for X (e.g., Fischer and Ravizza 1998). If I would resist eating sugary foods, knowing that my life was on the line, then I would be moderately reason-responsive and hence responsible for eating such foods under normal circumstances. However (and this has been used as an argument against this theory), it has widely counterintuitive implications to use these extreme conditions in discerning whether people are responsible in such situations. For example, if someone suffering from extreme agoraphobia never goes out even for a quick stroll, she is responsible for staying inside, assuming she would run out if her apartment caught fire (Mele 2006). But surely, this is not the case. She might be responsible for not seeking help or perhaps not using the techniques available to her, but she is not accountable for not going out; she cannot help it.

It has been argued that the theory has a wide range of plausible implications, which should give us reason to reconsider our intuitions regarding this case. Furthermore, people under certain conditions, even though they are responsible, are not blame- or praiseworthy (Fischer 2006). First, this argument is not that convincing, since the “counterintuitiveness” is so vast, and second, if we cannot blame or praise people, it is uncertain that all relevant aspects are left for our purposes. What we want to say is that when people are responsible for their ill health, then we have reason for holding them such. Whether or not this is the case concerning the addict is uncertain if we cannot take for granted that we have a *pro tanto* reason to hold addicts responsible when they are so.

will use a loosely defined idea that most of these notions have in common, namely, that if addicts to a substantial part have an impaired impulse control and because of it, in general fail to abstain from what they are addicted to, then these addicts are not (fully) responsible for continuing their addiction. This immediately sparks two further questions. The first is about whether addicts actually are constituted in this way, and second, whether they are derivatively responsible for their addiction in the sense that it can be traced back to an earlier point in time when they could avoid becoming addicts. I will say something about these questions in order.

The first question about to what extent it is true that addicts have an impaired impulse control so that their responsibility is mitigated is perhaps less straightforward than many people might think. Studies have shown that although addicts might have less impulse control than others, it is much higher than many people think. For example, when individual drug addicts are forced to think about whether it is worth continuing to take drugs or when they become parents or get a job that is incompatible with taking drugs, many of them quit (Elster 1999, Foddy and Savulescu 2010). Those who continue might be people who actually take great joy with or at least would suffer more without their drug use. These individuals are perceived by many as responsible.⁶² However, it does not mean that all addicts are fully responsible. There can be differences in both the individual level and the particular type of addiction. For example, it might be that people addicted to fat and sugar are excused for maintaining their addiction, while people who use heroin are responsible for continuing theirs. The point is that addiction is not always an excuse; hence, we cannot take it to be an unproblematic argument in favor of people in general being fully excused for ill health resulting from addiction. This position is not readily admitted in the literature on responsibility and health but is something we have to consider when discussing issues such as whether the obese are responsible for being such and whether they should be held responsible by getting lower priority in priority-setting contexts.

⁶²This is closely related to Harry Frankfurt's (1971) classic case of the "willing addict." If individuals wholeheartedly endorse their addiction, then they are responsible for what they do because of it. Many of us would judge people who do not take the pill to be responsible for it, and this is easily explained by Frankfurt's account. What matters to responsibility is not being able to do or act otherwise; it is that one's actions flow from one as a person.

The other question was about whether addicts are responsible for having become addicted. Even if the popular view on addicts and their non-existent impulse control is true and the empirical views described above are false, it is not necessarily the case that many of those who are suffering from ill health because of their addiction are excused. It seems reasonable to think that we are accountable for behavioral patterns we are responsible for developing, even though we cannot change these today. This process is often called “tracing” (Vargas 2005). If we can trace our addiction back to a point where we, for example, knew we would develop an addiction if we continued the way we did but did not care enough to quit, then we are responsible for being addicts now. This case might be true about certain kinds of addiction, such as alcohol and drugs. Many people know that there is a genetic and/or hereditary component to these forms of addiction. Thus, if someone close to them was an alcoholic, and they were old enough to recognize that this situation put them at risk of becoming alcoholics themselves, they would be responsible for their ill health resulting from alcoholism if they chose to start drinking in a way where this was likely to happen.⁶³ Conversely, people may not be responsible in the same way for their fat and sugar addiction. Even today it is a controversial issue whether these addictions actually are such in any proper sense, more so when these forms of addiction developed 15 years or more ago. But these are all open questions and something we have to examine when discussing whether the obese are responsible for their ill health.

To be responsible, we also have to fulfill the epistemic condition; we have to know or blamelessly not know about the relevant facts at hand. Regarding obesity, many people have said that the obese were either unaware of how to eat properly when gaining weight or now lack the knowledge of how to lose weight. However, we might think that they should have known and should know how to eat; hence they are not excused. The lack of awareness of facts we should otherwise have is usually called “culpable ignorance,” whose conditions and implications are widely discussed among philosophers (see e.g., Smith 1983, Zimmerman 1997, Rosen 2004, FitzPatrick 2008, Levy 2009, and Sher 2009). This discussion is often focused on “normative ignorance,”

⁶³Of course, they would also have to be aware of the fact that addiction might lead to ill health, which is just something I assume many people know, at least concerning alcohol and drugs.

i.e., not knowing the normative reasons, but there should be no problem applying what is said here to our discussion about responsibility and health. In fact, as we learned above, we might believe that we have normative reasons to be in good health and that people are culpably or blamelessly unaware of these. Even if we do not believe that we have normative reasons to take care of our health, the discussion about normative ignorance seems relevant for our purposes.

Roughly speaking, there are two views, one demanding and one less so. According to both, we need an opportunity to transcend our ignorance. How robust this opportunity should be is a bit unclear, but it is equally such on both sides. The difference between the two positions is that according to the more demanding view, we have to apply an idea of tracing, which is not the case with the more lenient idea. Begin with the more demanding notion (“the strong view”). To be culpably ignorant, we had to know previously that this ignorance would occur and that we had decisive reason and an opportunity to acquire the necessary knowledge but failed to do so (Rosen 2004: 306–308), i.e., we had to act at some earlier point against our better judgment.⁶⁴ For example, if someone with an unhealthy lifestyle at some point thinks that he or she should search the Internet on how to lead a healthy life based on the World Health Organization or the national recommendations, but he or she fails to do so because of weakness of will, then he or she is culpably ignorant in not learning about the facts. At least this holds true if we assume that he or she would have found correct online information about health and wellness.

The less demanding notion (“the weak view”) states that no such event of *acrasia* needs to be involved for us to be culpably ignorant (cf. FitzPatrick 2008: 608–610). Instead, the idea is that we are culpable for being ignorant when we are predisposed a certain way, irrespective of how this has come about. Perhaps more accurately, we are at fault for not knowing when culpability is a result of our own vices such as “overconfidence, arrogance, dismissiveness, laziness, dogmatism, incuriosity, self-indulgence, contempt, and so on” (FitzPatrick 2008: 609). The point is that the person who leads an unhealthy life does not have to have a certain moment where he or she thought that he or she ought to access information, but in the end did not because of

⁶⁴To act in this way is also called *acrasia* and is generally not perceived as an excusing condition.

acrasia. Rather, if one is lazy and self-indulgent, it seems quite plausible that one will never find oneself thinking this way. These thoughts just “bounce off,” so to speak, before they are considered seriously. Certainly the information has to be available so that you could have an opportunity to access it if you want to. But this sort of opportunity is much less demanding than the opportunity we need according to the strong view.

Cases in which probably most of us feel the strongest intuitive pull toward the strong view are when the *deliberation* resulting in a morally wrong act is described in some detail, it is clear that the source of the faulty information ultimately lies outside the agent’s control, and there is an explicit question about what the agent in this situation should have done instead (see for example Rosen 2004: 305f, Levy 2009: 737). Take, for example, the case with Bill, who decides to lie to his wife even though there were decisive moral reasons not to do so (Rosen 2004: 305f). Bill thought in his childhood that he should, when in conflict, favor prudential over moral reasons. It is not that moral reasons do not count, it is just that they do not have the same weight. So in this case where he knew that he would suffer a lot if he told her the truth, he, after some deliberation, came to the conclusion that he ought to lie. From his point of view, this was the most reasonable thing to do. Now, we might ask, what should he have done instead? Should he have told the truth? There were decisive moral reasons to do so, which means that he should have done it to do what is right. But was he responsible for not doing what is right? Of course, he could have told the truth against his better judgment. But is that really a reasonable demand? It would have been another matter if Bill at some point in time thought that this notion must be wrong, that we should give more weight to moral reasons than prudential ones. But as it happens, this is not anything that occurs to Bill, suggesting that he is not culpable for being ignorant and hence is not culpably ignorant when lying to his wife. He is just ignorant of the fact that this is the wrong thing to do and hence, is not responsible for it.

Cases in which probably most of us instead feel the strongest intuitive pull toward the weak view are when the agent’s action and its consequences are thoroughly described, the action is morally wrong and the consequences for others are negative, and the agent’s general abilities and opportunities are in

the foreground (see for example FitzPatrick 2008: 601–613, Sher 2009: 23–41). Take for example the case with Potter, who is a “reprehensibly ruthless” and powerful businessman (cf. FitzPatrick 2009: 599–609). He exploits the vulnerable position of the poor living in his housing projects to make small profits, profits he essentially does not need. This forces many people from their homes, families containing children and elderly, because the rent is demanded right away every month with no delay and no exceptions. People who have been his tenants for centuries but who now, on one occasion, cannot pay the rent in time are convicted because they had to pay for medicine, did not get their paychecks in time, etc., are cast out on the street. Potter also tries to keep the costs for maintenance as low as possible, this without telling his tenants. This puts their health at severe risk because mold infestation and other related problems are likely to occur. The reason why Potter does all this, why he is “sticking it to the poor,” is that he holds the false normative views. He learned from his father that it is only sissies and communists who care about the plight of the poor, and these views are something he basically agrees with now as an adult. Of course, they have not been unchallenged. Potter is well educated, and when he studied at the university, he met a lot of people with other views. The same goes for people from his community. Some people there share his views, while others radically oppose them. However, even though this is true, that his view has not gone unchallenged, he has not cared enough about the plight of the poor, so he has never thought hard about it. It is not that he has no ability to grasp moral reasons or to think critically about things. Far from it. He has just chosen not to apply these abilities on this case.

If I am correct, many of us have two different sets of intuitions when being confronted with the two cases. When we look at Bill’s case, we seem to think that he is excused, while when we look at Potter’s case, we seem to think that he is not. There is a discussion about what we should believe is relevant, and hence, how we should depict these cases (Sher 2009, Levy 2009: 734–741). I am not going to take a stand on the issue here.

Whichever view on culpable ignorance is correct, it seems that we might be responsible for at least some instances of ill health and injuries. For example, if someone living in an industrialized country has remained unaware of the hazards of smoking, then it seems quite possible that he or she, at some point in

time, has judged that he or she should seek information about it but failed to do so and is thus responsible.⁶⁵ Of course, in many countries, labels on cigarette packs warn and sometimes use graphic images about the dangers of smoking. However, this is not true for all countries, with the United States (US) as a prime example. Nonetheless, even in the US, it is common knowledge that “smoking causes cancer.” This fact suggests that most people are aware of this information. Similar lines of reasoning can perhaps be applied to other unhealthy behavioral patterns such as eating too much fatty food, exercising too little, and so on. The majority of people know that this sort of behavior incurs a high risk of poor health eventually, despite the absence of any warning label on the ordinary McDonalds meal. However, there might be many other forms of relevant ignorance, such as that one can become addicted to sugar, relatively few succeed in permanently losing weight after becoming obese, etc. Therefore, when discussing whether people are excused for their ill health in general and obesity in particular on account of ignorance, we have to examine whether they are actually ignorant and also whether they have had the capabilities and opportunities to obtain correct information. If they have, then it does not matter whether they are uninformed or not; they are still responsible. Furthermore, if we argue that people are not responsible for their ill health, we should use the weak criteria for culpable ignorance, and vice versa, to make it as challenging as possible for our own hypothesis.

There are more matters to cite when determining whether people are responsible for their health.⁶⁶ Nonetheless, these two objections, one from addiction and the other from ignorance, are among the important ones. Now that we have learned more about how these objections work, we know better how to approach the debate on whether the obese are responsible for being so. Moreover, if we draw the conclusion that the obese do not fulfill these two criteria, being addicts in a way that undermines responsibility (and more generally having a too impaired impulse control to be responsible) and being

⁶⁵ Of course, I might be wrong about this, but the only thing I want to do here is to give an illustration of how the discussion might go when examining whether people are excused because they were ignorant.

⁶⁶For an overview of the debate, see e.g., Sharkey and Gilliam 2010.

nonculpably ignorant, then we do not need to consider other potentially excusing conditions as well.

With this being said, the responsibilitarian criterion will be applied to the case of bariatric surgery. In the discussion on this topic, especially among ordinary people, the question about whether the obese are responsible has been and continues being widely discussed. We are also moving in the direction of having conditional welfare institutions where we are held responsible for things that entail our responsibility. This case is true today in certain countries, such as the Netherlands; others, such as Great Britain, might be moving in the same direction (Schmidt 2009). Thus, investigating whether people are responsible for being obese is a highly relevant endeavor. To do so I will, of course, determine whether people are misinformed and, if so, whether they are culpably ignorant and also whether they suffer from addiction and/or other impulse control impairments.

3.3.3 Utilitarianism

Utilitarianism is traditionally one of the most commonly held views among people involved in bioethics, healthcare ethics, and so on (Holland 2007, Beauchamp and Childress 2008), and it is easy to see why. People in the business of helping people who are injured and who suffer from ill health are interested in how they can make the situation better for their patients. In terms of priority setting, for instance, the natural question is what treatment would do the most good. Since professionals in health and healthcare institutions are among those steering the debate in bioethics, etc., the (traditionally) utilitarian focus is easy to understand.

Classic utilitarianism, i.e., hedonistic act utilitarianism, is the view that an act is right if it produces something of a more positive value, over what has negative value, than any alternative, combined with the idea that what has positive final value, a value in its own, is pleasure, and what has negative final value is pain.⁶⁷ Except pleasure and pain, a broad range of suggestions on val-

⁶⁷ There are different ways in which we can spell this out in more detail. For example, according to Tännsjö, we should define the act utilitarian criterion of rightness in the following way:

ues has been proposed, from those that are strongly related to the individual's quality of life, such as having one's intrinsic preferences fulfilled, to other things that are of objective value, such as having friends and an interesting job (see e.g., Brülde 1998).⁶⁸ The other classic version is rule-utilitarianism, which states that an act is right if, applied as a rule, it would produce more of what has final value than any other alternative (Kagan 1998).⁶⁹ Even though act utilitarianism and rule-utilitarianism are very similar, they might have different sorts of implications. For instance, putting my health at risk just because it is fun might be right, according to act utilitarianism, because it would, we may assume, create a larger total sum of well-being. If I injure myself, I can be healed; it would not cost anything extra for society as a whole, and I would gain a lot if I succeed. On the other hand, it might be wrong to subject one's health to unnecessary risk, according to rule-utilitarianism, because if people generally were to act in this way, it might result in significant costs for society, since many people would suffer from diseases and would be in need of healthcare.

However, many philosophers think that rule- and act utilitarianism coincide in many instances. There are two reasons for this. First, act utilitarianism has more than one level of analysis (sometimes called "two-level utilitarianism" [e.g., Hare 1981] or "sophisticated utilitarianism"). Expressed differently, the criterion of rightness might be different from the decision procedure.⁷⁰ For example, when a doctor thinks about whether he or she should hold his or her

A particular (concrete) action is right if, and only if, in the situation, there was nothing the agent could have done instead, such that, had the agent done it, the world, on the whole, would have been better.

According to the same theory, an action is wrong, if and only if, it is not right. And an action is obligatory if, and only if, had the agent acted in any way differently, the world on the whole, would have been worse. (Tännsjö 1998: 31)

⁶⁸Of course, it has also been argued that, for instance, equal distributions, organic wholes, and other things separate from what is good for the individual are of final value. I discuss some of these views under the other headings (egalitarianism, responsibitarianism, and autonomy).

⁶⁹There are many other forms of consequentialism, such as virtue (Driver 2001) and motive (Adams 1976). See also Kagan 1998 for an overview.

⁷⁰For a thorough discussion about the difference between utilitarian decision procedures and criterion of rightness, see Sidgwick 1907, Bales 1968, Hare 1981, Parfit 1984, Tännsjö 1998.

patient responsible for his reckless behavior, it might be true that it would be the right thing to do in this case, because it would maximize the total sum of well-being. However, possessing a character trait that makes it possible to hold someone responsible who is vulnerable and in need of care would, on the whole, we might assume, lead to worse consequences than being more lenient. If it is true that having a more forgiving character would yield better consequences, then according to act utilitarianism, we have reason to adopt such an attribute.⁷¹ Similarly, it is sometimes assumed that (act) utilitarians must defend institutional practices that maximize utility each time they make a decision.⁷² Moreover, it seems that many have thought that responsibility-sensitive institutions need to be justified on the basis of responsibility and desert (Walker 2010: 200f), but this is not so. Act utilitarians might have utilitarian reasons for building institutions that do not *try* to maximize the expected total sum of well-being when the latter attempts to decide what to do. This is because this action (the creation of the institution that tries to maximize the total sum of well-being) would not, we might assume, maximize the total sum of well-being.

The second reason why act and rule-utilitarianism in practice often coincide is that the main contenders of rule-utilitarianism take the actual consequences of people's actions into account when judging whether they are morally right. For example, according to Brad Hooker (2000), the most viable rule-utilitarian approach is one which "evaluates rules on the basis of the expected value of their acceptance by the overwhelming majority" (Hooker 2000). But what might these rules be? According to Hooker, they are the ordinary ones of thou shalt not kill, steal, and so on. But how do we come to the conclusion that these things are wrong? By thoroughly examining what

⁷¹Adopting this disposition leads to the wrong act in our particular case. The right act would have been to hold the individual responsible instead of not doing so. These sorts of acts are often called "blameless wrongdoing," because if what we do is wrong according to the act utilitarian criterion, it is not something for which we should be blamed, since it is a good thing overall if we are so disposed.

⁷²It seems quite obvious that utilitarianism can incorporate the idea that we should set our priorities according to more measures than some sort of utility-maximizing criterion, irrespective if there are other master values. If people feel that a certain way to prioritize in healthcare is gravely unjust, then they might be less inclined to pay for the healthcare given, and their trust for the healthcare institutions might decrease. Both of these things might in turn lead to a range of negative effects, such as worse health of the population. Alas, there might be utilitarian reasons to create institutions that are not only using utilitarian decision-making procedures.

the consequences would be if we did not accept these rules or perhaps accepted an opposite set of rules. The first step to do this, to see what rules might be, is to investigate what the actual consequences are from murdering, stealing, etc., so that we have something to go on. The next step is to see what different rules of thumb or different institutional rules together yield. For example, it might be the case that a rule stating “when there are poor, force the rich to help” in itself has extremely beneficial effects because of the law of diminishing marginal utility. But it might still be true that having a framework of letting the rich “keep their spoils” would have better overall consequences, because it makes people work harder, which in turn creates more for all. This way of trying to come up with what rules to follow and what institutions to build is, of course, very similar to the way sophisticated act utilitarians used above, which seems to suggest that rule- and act utilitarianism, in practice, may almost be interchangeable. What I have done is follow a common practice in bioethics, etc., of understanding utilitarianism as a view that commands us to try determining the possible consequences of creating certain institutional practices and performing certain types of actions, and then choose those with the most beneficial consequences.⁷³

When applying utilitarianism to different cases, it is important to remember that all consequences matter. One way of doing this is to divide the question about the effects into direct and indirect types. For example, it is not only the direct effects from bariatric surgery that count when deliberating whether to prioritize the procedure (i.e., the effects will be on those being treated, the

⁷³That we might not have reason to believe we can hold any justified views about these matters is a different question. For an argument against us being justified, see Gren 2004. There are also many other questions about how we should understand the utilitarian criterion of rightness, which I am not going to discuss here, such as whether it should be understood as subjective or objective, maximizing or average, actualist or contrafactualist, etc. (For a recent overview on different forms of consequentialism, see Driver 2012). Even though these views are interesting, it would take us too far afield to discuss them all.

However, one distinction might be relevant to us, the one between “average” and “maximizing” utilitarianism. Take the case of in vitro fertilization (IVF); according to the maximizing view, it is good to add another creature to the world if that creature leads a good life, since that adds value to the world. From the average view, this is not true if the creature does not have a higher degree of well-being than the mean. Since the contribution of value to the universe will be much greater if the maximization criterion is correct, we will have stronger reason to give IVF higher priority than if the average view is correct. However, it seems to me that the average view has been decisively rebutted (Parfit 1984: 420–422). Furthermore, it is not entirely clear which effect, if any, it would have on the debates I am interested in.

cost in relation to these effects, and so on), but also the indirect effects, such as whether third parties are being affected by starting to believe that they do not need to eat well and exercise because they can always undergo surgery. Even though this distinction is a bit vague, I have found it useful and have therefore used it when discussing obesity surgery.

In trying to answer questions about priority setting in healthcare, as we have seen, utilitarians do not necessarily have to use a decision-making model where they attempt to find out what treatment or drug will produce the best overall consequences. This is because such a procedure might not create the best possible outcomes. However, according to many utilitarians, when setting priorities in healthcare, we need to be able to state the consequences of prioritizing bariatric surgery, for instance. At least we need to know the effects of the treatment and how many resources are needed. This need not be the only variable to consider; other factors, such as the level of need and the degree of responsibility, might matter as well. Nonetheless, cost-benefit analysis should be one of the tasks we do when setting priorities.

There are many ways to measure the success of a treatment or drug. One of the most widely used health measures today is quality-adjusted life years (QALY).⁷⁴ The measurement was first developed in the sixties and seventies as a response to an increasing need for prioritizing in healthcare. The idea behind the measurement is to make it possible for us to find the cheapest possible treatments that guarantee the best possible effects in terms of good health (Nord 1999: 18–23). To do this requires some way of comparing “apples and oranges,” i.e., different forms of treatments for various types of injuries and ill health, and to this end, the notion of QALYs were devised.⁷⁵ In practice, healthcare institutions such as the British National Institute for Health and Care Excellence (NICE) established a cost ceiling for drugs and treatments per QALY, where one QALY is one year of perfect health, then estimates how much the drugs and treatments cost, phasing many of them out or not accepting all of those above the ceiling. For example, currently NICE has a ceiling of approximately 20,000 to 30,000£/QALY, which means that if a drug is more expensive than that in relation to the returns in terms of years of good health,

⁷⁴For a thorough discussion and critique, see Nord 1999.

⁷⁵There are three ways to measure these, see Nord 1999: 19.

the institute will perceive this as a strong reason not to offer the medicine to its patients.⁷⁶

It is easy to see how to connect the idea and use of QALY to the utilitarian criterion of rightness. Presumably, when discussing questions on priority setting, a utilitarian is generally interested in maximizing the total amount of well-being per resource unit.⁷⁷ When discussing priority setting in healthcare, it seems reasonable that the utilitarian should instead pay attention to, on average, maximizing the total amount of health per resource unit. This is because many types of ill health decrease well-being, and it might be a good idea to focus on ensuring good health vis-à-vis healthcare institutions, since it is the area of competence of health professionals. It might also be a good idea to impose a cost ceiling for treatment, etc., as NICE does, since it makes the priority-setting process more manageable in practice. Of course, the ceiling level is a difficult question, because it should be where it offers the optimal consequences, which are not easy to determine. However, current ceilings, assuming they are not justified, might still be useful, since if we maintain their limits, the drug or procedure we are discussing is at least not ineffective, compared to available alternatives.

As is to be expected, because utilitarian reasoning is controversial, the use of QALY in itself has been heavily criticized as is the use of ceilings for cost efficiency. For example, QALYs do not take into account who is getting the resources, only what the effects are of the resources spent. This situation implies that people who are in worse condition might not be prioritized over those in relatively better shape (because the latter might be easier to treat), which in turn has been perceived as gravely unjust. In many countries, this critique has led to supplementing the QALYs with other tools when setting priorities in healthcare. Principles such as that of the “greatest need” have been used, and other measurements such as disability-adjusted life years (DALY), health-adjusted life expectancy (HALE), etc., have been employed in parallel to compensate for the deficiencies in the QALY measurement.⁷⁸

⁷⁶<http://www.nice.org.uk/newsroom/features/measuringeffectivenessandcosteffectivenesstheqaly.jsp>

⁷⁷See e.g., the utilitarian Peter Singer's (2009) article in *The New York Times*.

⁷⁸It should also be noted that denying people access to drugs and treatments because they are too expensive has also created a lot of controversy. For example, there was a large outcry when women with ovarian cancer were denied the new drug Avastin, mainly because it would cost about 150,000€/QALY and

The use of QALYs and ceilings, then, might be something utilitarians can condone or at least use when trying to make decisions in allocating resources. This is especially true because QALY is so commonly used in conducting evaluations that it might be difficult to find evaluations employing other measures. There are, of course, other non-health-related effects of different procedures and drugs relevant to utilitarians. For example, we may find that people become happier and more socially interactive from getting treatment, which is not necessarily something linked to better health. Instead, it could just be a positive side effect relevant to utilitarians, obviously because happiness and meaningful relationships contribute to beneficial outcomes. The point is that although estimating the cost of each QALY is a good method for determining one's action, which outcome will produce the most good when allocating resources and other data on the effects of the procedures and drugs on matters relevant to people's well-being are noteworthy as well.

In this regard, when discussing whether we should prioritize bariatric surgery, I will use the utilitarian consideration, emphasizing whether it is cost-effective but also its other possible effects relevant to utilitarians, both positive and negative. I will also apply the utilitarian reasoning when discussing what perspective policymakers should take on people's responsibility. It will entail a broader discussion, not focusing on individual procedures or drugs, which is why the subject about QALYs will be absent. Concerning healthcare personnel, I will use the placeholders of beneficence and non-maleficence (more on these views in section 3.3.6).

3.3.4 Egalitarianism

The third normative theory is that of egalitarianism. This roughly expresses the view that we have reason to distribute resources or the good equally and/or that we should put focus on those who are worse off. For example, many countries aim to at least give their citizens equal opportunities for good health. This goal might imply a range of regulations, from "no one should be denied access to basic healthcare" to "the determinants of health, i.e., the con-

hence be too expensive. However, it is difficult to know whether this involved the grounds on which the priorities were set or was simply the priorities made. (For the outcry over Avastin, see <http://www.telegraph.co.uk/health/healthnews/9839289/Thousands-of-women-with-advanced-ovarian-cancer-left-frustrated-after-Nice-denies-them-drug.html>).

tributing factors to good health, should be distributed equally” (e.g., Daniels 2008).

Egalitarianism, then, is a widely popular notion. It is also an extremely complex concept that has been developed much further than this basic intuitive illustration I just presented. The three most common forms of fundamental and broadly egalitarian considerations are those of sufficiency (e.g., Frankfurt 1987), priority (e.g., Parfit 1991), and equality (e.g., Temkin 1993).⁷⁹ Sufficiencyarians believe that we have a moral obligation to help people up to a certain threshold, but we are not required to do anything else for those who have reached or are above this threshold.⁸⁰ Priorityarians deem that it is more valuable to help people who are (absolutely) worse off, and equalitarians judge that equality of outcomes has value in its own right.⁸¹ This means that priorityarians have reason to make the best of even those who are in relatively better circumstances, all else being equal, if this is the only thing they could do. In contrast, assuming that these individuals are over the threshold, sufficiencyarians do not have any reason to make it better for those in the best positions; neither do the equalitarians, assuming that people who are in worse situations will stay the same.⁸² Conversely, these different egalitarian considerations pull in the same direction when we have a group under the threshold; we can help up to not too extreme costs and where the uplifting of the group makes an outcome more equal.

There are also different forms of “currency” of egalitarian justice, such as welfare (Arneson 1989), resources (Rawls 1971, Dworkin 1981, Daniels 2008), advantage (Cohen 1989), and capabilities (Nussbaum 1992, Sen 1992). According to the “welfarists,” utilitarians have the value part right; what we should equalize is well-being. Exactly what that is—happiness, pref-

⁷⁹These views are often combined with utilitarianism or some form of responsibility view, because they are often considered favorable for the egalitarian views and the responsibility and utilitarian views, respectively. For example, according to Derek Parfit (2011), the most problematic feature of utilitarianism can be solved if we combine it with some sort of egalitarian reasoning (Parfit 2011: 196–201). Moreover, according to Carl Knight (2009), the same is true regarding the responsibility considerations.

⁸⁰For a recent overview and critique of the sufficiency view, see Casal 2007.

⁸¹This is of course a huge simplification on my part. For example, not everyone is interested in outcomes; instead, according to them, what matters is that everyone has equal opportunities.

⁸²For a similar line of reasoning in the policy area of priority setting, see Alan Williams’ (1997) “Fair innings argument.”

erence, satisfaction, or fulfillment of some sort of objective list, can be discussed. However, whether the value can be found among these theories of value is not open for discussion.

According to the “resourceists,” we should not equalize welfare but external and/or internal goods, which can be used to attain whatever the individual wants, whatever her version of the good life is. This can be “primary goods” such as income, education, and health. Those believing in the idea about advantage take a middle ground between the resourceists and welfarists; welfare deficits matter in some cases even when they do not impact on our resources and vice versa.

Lastly, according to the capability view, we should equalize “real freedom,” i.e., people’s abilities, so that they can do whatever they want. This view is similar to that of the resourceists. Both are interested in real freedom and believe that it is of utmost importance that the individual has the choice of his or her own good, with the difference being that those advocating the capability approach only care about “stuff” as an instrument for distributing.⁸³ In practice, when discussing health and healthcare policies, these different views will often pull in the same direction. The reason is that helping those who suffer from ill health will help those with lower degrees of welfare, capabilities, resources, and advantages, in comparison with the general population. Of course, the scope of whom to help and the extent of help will differ, but when we are discussing what to do about people suffering from ill health in our actual and unequal societies, these different views tend to promote prevention through public institutions (e.g., infrastructure for distribution of clean drinking water to all, vaccination programs, etc.) and publicly funded healthcare, up to the point when the respective egalitarian endpoint has been reached, such as a reasonably equalized access to health resources in relation to a minimum level of health for all.

Egalitarianism is also either outcome- (Williams 1997, Persson 2006) or opportunity-based (Dworkin 2002, Daniels 2008). For example, according to

⁸³The egalitarian views are even more complex than what I have described here. There are also questions about among whom (Valentyne and Tungodden 2006), when (Temkin 1993), how we should measure inequality (Temkin 1993), etc. I have attempted a brief sketch of what the ordinary discussion is usually like.

Norman Daniels, for society to be just, everyone should have the same opportunity to pursue one's own personal good, meaning the same chance to fulfill one's own life plan.⁸⁴ Of course, different life plans demand different things, but one often needs at least decent health to have a reasonable chance of reaching one's target. This is true whether one wants a career as a high-powered business lawyer or a homemaker. In other words, health is in this way special, compared with other goods. Decent health is the basis for having one's share in a "normal opportunity range," to be able to do the same things as others do in society.

Now, since we can to some extent distribute both health (so to speak) and healthcare, we should do so, according to Daniels, with an eye on equality. Equally distributing determinants of health and healthcare implies for instance that healthcare should be universal and publicly funded, and preventive work should be performed in all societal sectors. After this distribution has been made, when everyone has the same opportunity range, it does not matter from the view of justice how people fare. It might be a bad thing if they use their options poorly, but we do not have justice-based reasons to amend the consequences of these choices. (At least not more than supplying them with healthcare and other things relevant to maintaining the same opportunity range).

However, according to the outcome-based views, we have justice-based reasons for doing so. It is unfair if people fare differently because of their choices or other things; hence we have reason for trying to compensate so that those who are worse off become better off and come closer to the others.⁸⁵ Nevertheless, in practice, with regard to health and health policies, the opportunity and outcome views often coincide. People often have different states of health and thereby different levels of welfare because of not having an equal opportunity to health in the robust way Daniels described. And lack of opportunities many times is brought about through access to concrete resources, where

⁸⁴For a detailed defense of Daniels' view, see his book, *Just Health* (2008), especially chapters 1–5.

⁸⁵Of course, both the outcome and opportunity accounts can be combined with other views, such as the responsibility view, so that some welfare deficits, for instance, do not count even to the outcome egalitarian. It should also be noted that we can reach equality in more ways than lifting up those who are worst off; we can make it worse for those who are better off while holding the worse off constant (i.e., "leveling down").

health is one. That is, worse health in practice weakens one's opportunities for attaining health in the future.

To sum up, egalitarianism is a complex cluster of notions. However, since my aim is to discuss policy in the real world by applying the egalitarian views on the problem of whether we should hold people responsible for their injuries or ill health, we do not have to worry that much about the intricacies of egalitarianism after all. This position is true not only with the outcome/opportunity distinction, but also, as we have seen, with the other distinctions, at least with the case of a group under the threshold being worse off than others, whom we can lift up to a reasonable cost.

The conclusion is, then, that when we consider whether to hold people responsible on egalitarian grounds, we can pick one theory and test it on the case at hand. After that, we have at least some reason to believe that we have an answer that suffices for all the different notions. Of course, this is not a guarantee; we probably need to spell out the case in detail and apply the different egalitarian views to be certain. Nonetheless, since these varying egalitarian notions often point in the same direction, we will take a great leap, answering the question about whether to hold people responsible and create responsibility-sensitive institutions. Now, relating the discussion about egalitarianism to questions 3 and 4, I will use the outcome-based welfare view when discussing whether policymakers should perceive people as responsible for their ill health. As for whether we should prioritize bariatric surgery, I will use the placeholder view of the greatest need. (The principle of the greatest need will be discussed more in section 3.3.6).

3.3.5 Autonomy views

Thus far we have learned that many factors are relevant when thinking about holding people responsible for their ill health and distributing responsibilities for people's health. For example, we might think that society should take responsibility for people's ill health through legislation (i.e., banning narcotics and imposing heavy taxes on cigarettes and alcohol) and enacting publicly funded healthcare to which all of society's members should contribute, because this line of action would yield the best consequences. Expressed differently, distributing much responsibility to society for our health produces the

best possible outcome, and if utilitarianism is true, we are justified in doing so. However, a long tradition states that “the right precedes the good,” where it is generally thought that even though it might be good for people to be healthy, and even though such a handling of people may further the general level of well-being and/or egalitarian patterns of health distribution, it is nonetheless not allowed for society to use people as mere means for good ends (Kant 1785, Rawls 1971, Nozick 1974). This may regard both the way in which policies of the kind just mentioned “nudge” or even force people to make certain health choices without their explicit authorization and the fact that these policies require money that is extracted as tax from citizens, whether or not they like it.

Many people believe, “I am entitled to do what I wish with my body as long as I do not hurt anyone else.” For example, if I visit my physician, who thinks I would be better off losing some weight, he or she cannot force me to start exercising, take weight-reductive drugs, or have surgery. He or she is only allowed to try to convince me why I should follow his or her advice and how it could be done, and this is how it should be. We would not want our physicians to do things to us of which we do not approve. Of course, there are many difficult cases when the individual does not want treatment but clearly it would be beneficial for him or her, and we are inclined to think that overriding the individual’s decision would be allowed. Cases of suicidal young people and mentally ill adults refusing to take prescriptive drugs are just two examples.⁸⁶ However, if nothing else, as a rule of thumb, many of us think it is paramount for physicians to obtain our consent before they start the treatment (Faden and Beauchamp 1986, Beauchamp and Childress 2008).

We might also claim that we have similar rights to what we rightfully own as well. Assume that I have worked hard for a long time and saved all my money for a new car. It is not that I need a new car that badly, but I would really like one in order to arrive more quickly and comfortably at work and to drive to the countryside during the weekends, hiking and just enjoying myself. After a year of living on bread and water, working really hard, I finally have enough money to buy the car. It is not the fanciest, but also not the shabbiest. However, as it happens, my neighbors also need a car, and they

⁸⁶For a thorough discussion about these sorts of cases, see Tännsjö 1998.

would have more use of it than I, i.e., if they had the car instead of me, the universe would be a better place, since it would contain more of what is of worth. This case would even be true if we factored in my disappointment at getting my car taken from me. If they had a car, they would be able to get better jobs and drive their three kids to their different activities, which would generate much more good than my trips to the countryside. However, many people would say, even though it would be better if society took my car and gave it to the neighboring family, society is still not allowed to do so. The family could ask if they could borrow my car, and society might be permitted to try to persuade people such as me to think about the less fortunate. We may even agree that I indeed have a moral duty of care to give away my car or sell it very cheaply to my neighbor. Nonetheless, it is not allowed for others without my consent to redistribute my car to the needier where it would do more good. If this is true, something similar could perhaps be said about taking my money through taxes and using it for the greater good, which, for instance, might mean funding healthcare. No one says it would not be a good thing if others became better off, but as with the car, it is still wrong for others to use my (hard-earned) money against my will, at least unless it is necessary for upholding the conditions of a just social contract.

There are many different ways to defend these claims; one of the most prominent ones is that of libertarianism. According to the standard libertarian view, formulated by Robert Nozick in his *Anarchy, State, and Utopia* (1974), we have *absolute natural rights* to our own bodies (self-ownership) and the holdings to which we are entitled (entitlement theory) (Nozick 1974: 150–182). It means that others are not allowed to intervene in our lives, forcing us to do or not do different things, as long as we do not infringe on someone else's natural rights. Furthermore, others are not allowed to take our belongings without our consent, at least if they are not necessary for society to uphold and defend said rights—in the libertarian view, basically the criminal and civil legal systems and the military. If I do not want to exercise as my physician recommends, then he or she is not allowed to force me, even if it would be good for me. Similarly, even if my car would do more good in the hands of my neighbors, they, or anyone else for that matter, are not allowed to take my

car against my will. If I want to leave it in the garage, costing me money without using it, I may do so; no one else has the right to force me otherwise.

The standard libertarian view seems to give us the right answer to the two preceding examples. No one, not the state, the church, nor our family or friends, has the right to force us into doing something we do not want and take something we rightfully own. Stated another way, there are “side constraints” on what we can do to people and with their belongings (Nozick 1974: 28–35). This libertarian stance is strict about respecting rights; it does not allow us to “maximize rights,” fulfilling some people’s rights by infringing those of others or something similar. In cases where we would be able to minimize the amount of right violations by violating someone’s right, we are still restricted from doing so. This view has also been used in many contexts related to this dissertation, such as that of bioethics in general (Engelhardt 1996) and our obligation to finance other people’s healthcare in particular (Engelhardt 1996, Narveson 2006). Even though on the surface, we might judge this to be a plausible view, it seems to suffer from several insurmountable difficulties, while other plausible considerations capture the intuitive appeal of libertarianism, all of which suggest that we should give it up. Before I explain why this is so, it should be noted that all these arguments against libertarianism that I will present have been developed and defended by others, so I do not by any means claim to be original.

Begin with the notion of entitlement. According to Nozick:

If the world were wholly just, the following definition would exhaustively cover the subject of justice in holdings:

- (a) A person who acquires a holding in accordance with the principle of justice in acquisition is entitled to that holding.
- (b) A person who acquires a holding in accordance with the principle of justice in transfer, from someone else entitled to the holding, is entitled to the holding.
- (c) No one is entitled to a holding except by (repeated) applications of (a) and (b).

The complete principle of distributive justice would say simply that a distribution is just if everyone is entitled to the holdings they possess under the distribution. (Nozick 1974: 151)

Both (a) and (b) are quite unclear, and their interpretation is a matter of debate (Wolff 1991). Nonetheless, the rough idea is that we acquire a holding in accordance with the principle of justice when it is not formerly owned and we do not make other people worse off by acquiring it (the Lockean proviso). Since the world from the beginning was not owned, everything was “up for grabs,” at least as long as the Lockean proviso was met. Therefore, if we all lived in the desert with only one waterhole that all needed to use, then it would be an act of injustice if one of us tried to make it his or her personal property, because everyone else would be under worse conditions. But if there was one waterhole for each of us, and one was all we needed, then there would be no injustice if someone or more claimed one of these holes for their own. Furthermore, through our acquisitions, we are not making anyone else worse off if they could have avoided being such through some action(s) of their own. If there were a risk of the water being poisonous, and I had ways of avoiding this, then if only I chose to take these safety measures, I would not have made anyone else’s situation worse when the catastrophe came, making all water undrinkable except mine. Hence, I can do as I please with my water, and no one else is allowed to force me otherwise.

I think we can already recognize where one problem lies with the entitlement theory. When we discussed the negligent and negligent victim objections, the idea that we do not have any justice-based reason to help those who are responsible for being injured, we saw that this is generally considered a strong objection against the view. Now, this challenge can be raised against the entitlement view as well. Consider the case of the desert where the water has been poisoned. Assume that the risk of getting one’s waterhole poisoned was very low and that people in the community did not even know of the risk. Assume also that only one hole was more than enough to give everyone access to water. Since most people in the community would share if the situation demanded, the risk of all waterholes being affected was extremely low, and because they could do other valuable activities with their money, everyone but one did not invest in protection for their water. Sadly enough, the catastrophe struck, and everyone except one got their water ruined. However, since this hole’s owner relished seeing other people suffer and die, she refused to give the

others any water. In the end, all in the community died but the sadistic and cautious owner. Is this really a just outcome? Certainly, it might be that the others in the community had an obligation to compensate the sadist for her spent resources and perhaps also some more for drinking her water. However, it seems too harsh for them to die just because they had chosen not to take every precaution against a very small risk. Accordingly, the entitlement theory, as it stands, has very counterintuitive implications, which, of course, is a good argument against it.⁸⁷

The entitlement theory is widely counterintuitive and hence seems implausible. Does the notion of self-ownership fare any better? Unsurprisingly, I do not think so. First of all, this theory has counterintuitive implications as well. Similar to the case with the waterholes, there are many instances where we think we are allowed to force individuals to do things for the good of others against their will.⁸⁸ If I can save my child's life by causing someone some sort of mild and nonlasting pain, and I only had such recourse because the person did not want to do something of insignificant cost or trouble to himself for my child to live, many of us would say I would be justified in inflicting the

⁸⁷Another problem concerns the Lockean proviso (Cohen 1986, 1995; Wolff 1991). The idea seems intuitive enough but is probably incoherent. We are not allowed to acquire holdings so that other people become worse off. I am not allowed to acquire the only waterhole in the desert because that would make everyone else much worse off. The baseline here is then when the resource is in general use, but why should we use this baseline? Take another case of two individuals who are both living off an acre of land (For a more detailed discussion, see Cohen 1995: 79–84, 90f.). Assume that A acquires the land but gives B the chance of employment on his land. B feels that she has to take the job because there are no other alternatives available. Since A has some organizational skills, both are better off after A's acquisition. A is much better off, and B is somewhat better off, suggesting that the Lockean proviso is fulfilled. However, assume now that if B had acquired the land, things would have become much better. B would have been much better off, and so would A. This implies that B, in the first scenario, is actually worse off after A has acquired the land; it would be much better for her if she had been in control of the land after all. However, if we were to accept this comparative baseline instead, where we considered whether everyone could have had it better under some other distribution of holdings, our acquisitions would never be just. The reason is that someone's circumstances would always be better if someone acquired the land other than the one who actually did. Since every individual is of absolute importance, we cannot justify holdings on the basis of the majority's improved circumstances under such and such conditions. This issue, that holdings would never be just, given some other baseline than the one Nozick suggests, might be what spurred him to propose this idea from the beginning. Nonetheless, for the baseline to be justified, he has to give some justification for it, which he has not done and which hence counts against his entitlement theory.

⁸⁸Derek Parfit (2011) convincingly (to my mind at least) shows that it is not always wrong to treat an individual as a mere means (212–233). This line of reasoning is similar to mine when I try to show why we do not enjoy self-ownership in the strong sense Nozick believes to be the case.

pain to save my child's life. Nonetheless, according to Nozick's theory of self-ownership, I am not justified in doing so. It would definitely be good if my child was to survive the situation, and the person refusing assistance is certainly of horrible moral character, but justice demands that I do not save my child's life by imposing my pain-inflicting strategy on someone. Of course, this implication is absurd, which gives us reason to reject as false the self-ownership theory.⁸⁹

Finally, even if all these difficulties could be overcome, the standard libertarian view would not have much to say about many of the cases of interest to us.⁹⁰ Beyond the ban on force, libertarianism is silent on "what we owe to each other" and will provide no guidance to issues about rationing scarce resources within a scheme allowed by libertarianism. Assuming that all are entitled to their holdings, the only thing libertarianism indicates is what we are prohibited from doing, not what we ought to do. For example, it seems quite likely that in a libertarian society, we would have charities providing poor people with healthcare. At least some of these charities would probably have

⁸⁹Nozick says that when moral catastrophes are pending, we are allowed to violate people's natural rights. Similarly, if our own life is at grave risk, such as if a person falls toward us when we are sitting on the bottom of a well, we are allowed to kill him with a laser gun if that is the only way for us to survive. The problem with these suggestions, besides coming out of the blue, is that it becomes difficult to know where to draw the line and to what extent Nozick's theory differs from the more moderate views presented below.

⁹⁰The reasons for this are twofold. First, we do not know, and with today's knowledge cannot know, who is entitled to what. For a person to be entitled to her holdings, she had to come into their possession in a way that fulfilled either condition (a) about just acquisitions or (b) about just transfers. Now think about the sources of our holdings. They probably come from a wide range of areas such as our salaries, bank loans, investments, provisions made by the states in which we live, etc. For our holdings to be just and for us to be entitled to them, all our employers, banks, markets which give us returns on investments, and the government must have acquired their holdings in a just way. Already it seems clear that many of these agents have not attained these holdings in a way conducive to (a) and (b). However, even if their holdings had been created in a legitimate way, for their holdings to be just, those from which they (banks, etc.) obtained their holdings should in turn have gotten their holdings in a way conducive to (a) and (b). For my holding to be just, each of these steps, from today back to the first time someone acquired something, must have been made in accordance to (a) or (b). Otherwise, my holdings are unjust, and the historical error should be corrected. The problem is that no one knows when and from whom these wrongs had begun, because no one believes that nothing has gone wrong, which means that we do not know who is entitled to what. We cannot even make educated guesses, because past wrongdoings are so great and are long since forgotten, implying that we cannot even start to reason about what to do when discussing responsibility and health. Of course, we could wonder whether this is just a practical problem for us when making decisions or whether this problem counts against the theory itself. I will not discuss this issue here.

an open mandate on whom and what to treat, and these beneficiaries would most definitely have scarce resources. There would be more need for healthcare than charities would be able to satisfy. Given their open mandate, a question will arise about whom and what to treat and in what order, and the libertarian theory has nothing to say to this. However, the libertarian view might have something to say about, for instance, paternalistic ways of holding people responsible, but these areas of application would be much slimmer. This is because we would be unable to answer the broader question of what to recommend to doctors regarding how they should perceive their patients, for instance. According to libertarianism, doctors do not have any obligations to their patients; if their patients do not like their doctors, they should go somewhere else. This means that we do not get any guidance on the question at hand. Of course, libertarianism might allow that a complementary ethical theory is brought in to provide proper guidance. However, that would be exactly that—another theory—and not libertarianism.

Thus, as it stands, the standard libertarian view is unclear, vastly counterintuitive, and hence not a plausible, fundamental normative consideration. Even if it were, it would not provide any answers to many of the questions relevant to this dissertation. However, it should be noted that many of the problems with Nozick's view on self-ownership and entitlement constitute the absolute character of these natural rights. It seems that a more plausible view is to argue that autonomy has a value on its own but that we do not always have decisive reason to promote it, because some sort of pluralism is true. For example, according to this "alternative view," it would be a good thing in itself to lead an autonomous life according to our own desires. However, if other great values were at stake, these would in some cases trump the value of autonomy. In these cases, such as that of saving of a child, where we could do much good with only a small infringement of the person's autonomy, we would be justified in doing so. Otherwise, in most other cases, where the infringements would be considerable and the benefits not as significant, the value of autonomy would stand. This is probably true even if (as some utilitarians do) one takes autonomy to be an instrumental value only. For example, the hedonistic act utilitarian Torbjörn Tännsjö argues against the use of coercive care on the grounds that violating people's autonomy would have negative consequences

(autonomy has a positive instrumental value) (Tännsjö 1998, see also Dworkin 1988).

Thus, when discussing whether we should perceive the skeptical arguments as viable, I will ask whether this might lead to actions and attitudes that violate their autonomy. This consideration is most often used when addressing questions at the “lower levels,” e.g., the relation between healthcare personnel and patients, which is why I will apply the autonomy consideration to the question of whether healthcare personnel should perceive the skeptical arguments as viable.

3.3.6 Greatest need, beneficence, non-maleficence

Aside from all the fundamental normative considerations, there are also several important derivative normative considerations in the debate, including those from beneficence (do good), non-maleficence (do no harm), and the principle of greatest need (give priority to the worst off).⁹¹ These principles are quite “hands on” in the sense that they tell us what to do in our line of work. We should not harm people, and we should try to do good and focus most of our efforts on those who are worst off, when we can do something that is not grossly cost-ineffective.

I will use the principles of beneficence and non-maleficence when discussing how health professionals should perceive their patients. This is simply because these principles are often used when discussing how people at this level should relate to their patients. These principles will also work as a placeholder for utilitarian reasoning. If we choose a strategy that creates the most good and the least harm, we will probably get one that creates the most utility. The principle of greatest need will be applied in the case of bariatric surgery, because it is commonly used in priority-setting contexts, and it (or a similar principle) is the law in many places (Lie and Sabik 2008). The principle of greatest need will also be a placeholder for different forms of egalitarian reasoning. After all, this principle makes us focus on people who are worse off and what we can do for them to have their lives improving closer to those who

⁹¹ The principle from greatest need is often understood as proposing something stronger, namely that the worst off should get absolute priority in the sense that we should put all our resources here (SOU 1995: 5). However, the problem with this principle is that it is vastly counterintuitive (Daniels and Sabin 2008).

are better off. In health-related contexts, we can be quite sure that this group will be under the threshold (sufficientarianism); they are worse off, and it will have more worth to lift them up (prioritarianism), hence creating less inequality (equalitarianism).

3.3.7 Summary and outlook

In summary, when answering the question of how policymakers should perceive people (as links in chains or not) I will use broadly utilitarian and outcome-based welfare egalitarian considerations. When addressing the question of how healthcare personnel should perceive people, I will apply the autonomy view, together with the considerations from beneficence and non-maleficence. Finally, in examining whether we should prioritize bariatric surgery, I will employ a broadly utilitarian consideration, together with responsibility and the principle of greatest need.

PAPER 3: THE RIGHT PERSPECTIVE ON RESPONSIBILITY FOR ILL HEALTH

In Paper 1 we introduced the “explanation hypothesis” and showed how it could explain our everyday responsibility judgments as well as the appeal of skeptical arguments such as those from luck and heteronomy. In Paper 2 we described how the explanation hypothesis could account for the recent findings and contradictions in experimental philosophy.

We have, then, in effect answered questions 1 and 2. In this paper I am going to try to answer question 3, arguing that those working in different positions with regard to health and healthcare should take different explanatory perspectives on responsibility and hence should differently perceive those who are in need of healthcare and healthcare prevention. People (such as physicians and nurses) who interact directly with patients should take an everyday perspective and judge patients (or others in need of preventive healthcare) to be significant explanations to their ill health and hence be prone to perceive them as responsible. Policymakers, who set up the institutions where the healthcare personnel and others work, should hold the opposite view, that the populace is not responsible. That is, people on the “shop floor” should focus on the motivational structures of the individuals, while policymakers should focus on what has caused these structures to come about.

As we saw in Paper 1, one implication for the debate on responsibility, given that the explanation hypothesis is correct, is that the validity of the skeptical arguments has to do with the normative reason at hand. For example, assuming that some sort of consequentialism is correct, and the best consequences come from judging people to be responsible, then we should take a perspective where we judge them to be responsible. The normative principles used on the different levels (the policy level and the shop floor) are, as we saw in section 3.3, somewhat different, even if they mostly are overlapping. On the policy level, the utilitarian and egalitarian notions are prevalent, and on the ground level, the notions of beneficence, respect for autonomy, and non-maleficence are instead the most common ones. I use these different notions to analyze what perspective to take on the different levels.

When deciding what perspective people should use on the different levels, we first have to predict what it is most probable that these people will do,

given that they take a certain perspective, then evaluate this behavior against the normative principle relevant to the different levels. For instance, if taking a skeptical perspective as a physician leads to a much greater probability of infringing on someone's autonomy in a disrespectful way, then the physician has a normative reason not to take this perspective (given, of course, that taking the other perspective would not lead to even greater autonomy violations).

Policymakers should take a skeptical perspective on people's responsibility for their ill health, because it will lead to a healthier population (the utilitarian notion) and the health will be more equally distributed (the egalitarian notion). There are many reasons why the populace will be healthier given that policymakers take a skeptical perspective where people are seen as "dependent variables." But the most important one is that when policymakers judge people to be exculpated from responsibility for their ill health, they become more prone to invest in different sorts of preventive measures. Since these sorts of measures are the most effective to create good health, policymakers have a utilitarian reason to take the skeptical perspective.

The main (outcome-based welfare) egalitarian reason why policymakers should take a skeptical perspective is somewhat similar. Vulnerable groups who suffer from worse health from the beginning are more likely to indulge in risky behavior which, on a population level, leads to an even greater gap between those who are worse off and those who are better off. If policymakers take a non-skeptical perspective, they are less likely to implement preventive measures and more likely to "punish the people" through higher insurance premiums, lower priority in priority-setting contexts, and so on. On the other hand, if policymakers take a skeptical perspective, they will be more likely to use the more effective measures there are, which makes the gap between the different groups a bit more narrow.

Those on the ground level, the healthcare personnel, should instead take a non-skeptical perspective. They should perceive their patients as responsible. This is because it will lead to better health, fewer autonomy violations, and less harm.

When the healthcare personnel meet with the patient, they can often do much more for the patient if they perceive her to be responsible. This is true at least in cases where the patient's actions are causally effective, she has a

normal degree of impulse control, and so on. When the healthcare personnel perceive their patients as active agents, making choices and so on, it makes them more prone to give their patients advice on how to act in certain situations to avoid health risks. If the healthcare personnel instead take a skeptical perspective, they will be more prone to perceive the patient as someone who has been unfortunate, with no way to turn things around. The same sort of reasoning can also explain why autonomy infringements might be more common if the personnel take a skeptical perspective: the patient is seen as less autonomous, which gives them less reason to not violate her autonomy. With regard to non-maleficence, the strongest argument in favor of using the skeptical perspective is that perceiving people to be responsible for their lifestyle-based diseases might do harm, because the healthcare personnel might subject their patients to reactive attitudes. However, I do not believe this problem to be insurmountable. Most healthcare personnel will understand that there is a time and place for everything and that scolding patients is not the proper thing to do. Furthermore, given the proper training, they should be able to see that perfect health is not all that matters; hence, there is often not much blame attached to lifestyle-based diseases.

To conclude, we should take different perspectives on different levels. This is because different normative principles are being used, and people on different levels are able to do different things. Physicians who meet with patients face-to-face can do one thing; policymakers who are involved in setting up the structure for society can do other things. Combined with some difference in normative reasons, such as those from autonomy, on the different levels, we get this interesting and diverging result.

PAPER 4: WHY BARIATRIC SURGERY SHOULD BE GIVEN HIGH PRIORITY: AN ARGUMENT FROM LAW AND MORALITY

In this paper I answer question 4, should we give bariatric surgery high priority, by arguing that we should give it high priority, given the most commonly used normative and legal principles described in section 3.3: the principles of greatest need, utility, and responsibility. I will discuss each of these in turn.

According to the principle of greatest need, one should focus on those with greatest need or who are at risk of becoming one of those with the greatest need. When distributing resources, this translates to giving those with the greatest need (i.e., those who are worst off) most of the resources.⁹² That the obese are among the worst off is quite easy to see. They are often at high risk of being affected by life-threatening diseases such as heart diseases and cancer. People who are obese also tend to suffer from an extremely low quality of life. They tend to also suffer from different diseases such as diabetes. All of these high risks would either go away or at least be mitigated if the obese would lose some of their weight (see for example Sjöström 2012). That the obese are among those who are in the greatest need is, then, not a very controversial claim to make.

The utility principle states that we should do what directly and indirectly maximizes utility. Here, we need to show that the procedure is effective in the sense that people who go through the surgery lose weight and that their utility, whatever that is, rises to a greater extent than with patients who undergo many other procedures. We also have to show that giving the performing of bariatric surgery high priority does not have negative effects that mitigate the positive ones. For example, it might be the case that if we give the procedure high priority, people in general would start eating more fatty and sugary foods and gain a lot of weight in a way that they would not if the surgery had low or

⁹² Some would argue that this is only the case when distributing resources this way has beneficial effects. That is, we should only give high priority to bariatric surgery if we can show that the obese person is in great need and the surgery would have positive effects. I will discuss the effects of bariatric surgery further down in the text, under utility.

no priority. If this is the case, then placing a high priority on bariatric surgery might do more harm than good.

According to the overview I did at the time when writing this paper, paying particular attention to what we can do for the obese, it seems that there is no evidence that alternative treatments have any lasting effects. For example, most of the classical therapies, such as giving advice on nutrition and exercise, only work in the short run but not in the long run. People are often back to their starting weight after two years. With bariatric surgery it is different. People lose a lot of weight in the beginning, and in general, they can retain their weight loss over the years. With regard to negative side effects, these are quite few and sufficiently far between. We also have no reasons to believe that placing a high priority on surgery would actually make the populace gain weight, and the medical side effects are so marginal that they do not have a great impact on the positive value of the treatment. Furthermore, it has also been found that people get a substantial “lift” from the procedure; they become healthier, happier, and so on from the weight loss. Bariatric surgery is also proven to be cost-effective in general, which means that the procedure keeps up well with procedures in other areas.

Last is the argument of responsibility. Often in popular debates, as well as in philosophical and policy debates, there are voices claiming that the obese are themselves responsible for being obese and hence only have themselves to blame. Based on these debates, it is implied that we should not have the same obligation to help the obese as we have to help others who are not responsible for their predicaments. Assuming that we in general are responsible for our behavior, that there are no general excuses such as those from determinism discussed in Paper 1, there are still many reasons to believe that people are not responsible for being obese. One quite straightforward argument is that people who are strongly motivated to lose weight in general do not succeed. There are a lot of reasons why this is so, of course, many of which do not excuse them. But it seems quite clear that we have good reasons to believe that the explanation for why the obese in general fail in losing weight has to do with exculpatory conditions that they are not responsible for, such as a severe loss in impulse control through addiction-like symptoms, being misinformed in a non-culpably ignorant way, and not being rational.

To see why sugar and fat addicts are excused while heroin addicts might not be, you could just think about how extremely difficult it is to avoid eating unhealthy foods when you have cultural norms pressing you to eat these foods and you are constantly reminded and manipulated to eat these foods by the companies that produce them. These things and many more make it much more difficult for a sugar addict to get control over her addiction than a heroin addict. There is also no common knowledge about how difficult it is to lose weight when you have become obese or that you actually can be a sugar or fat addict. This means that people, especially those who are adults today and who are eligible for surgery, have very seldom had access to information that might have helped them to not become obese in the first place. All of this comes down to the fact that the obese in general are excused and hence should not be held responsible.

In summary, the obese are a group of great need, not generally responsible for their being obese, and bariatric surgery is the only measure we have to create longstanding positive effects. For these reasons, we should place a high priority on bariatric surgery.

4. Results, implications, and further research

This dissertation spans a vast number of sub-areas within the general theme of responsibility for health, from theoretical discussions about the preconditions for responsibility to more practical considerations on whether we should prioritize bariatric surgery. A number of conclusions have been reached, which I will briefly summarize in this chapter. After discussing the results, I will also touch on their implications and avenues for further research.

4.1 RESULTS

In section 2, I posed four questions. The first was about the evidential value of our exculpating intuitions when confronting the skeptical arguments of heteronomy and moral luck. In Paper 1, I argued against such a position based on the *explanation hypothesis*. For us to judge someone to be responsible for some event, we have to perceive their motivational structure as a significant explanation of that event. We perceive people as excused when confronted with skeptical arguments, because these arguments make us adopt explanatory perspectives where the motivational structure falls into the explanatory background and hence becomes explanatorily insignificant. The motivational structures are perceived as just being links in chains. However, the explanatory perspectives that elicit skeptical intuitions are just one set among many; from other perspectives, people seem responsible. So we need some sort of argument in favor of making judgments from these skeptical perspectives instead of the other everyday perspectives.

The arguments in favor of using a certain set of perspectives could be both metaethical and/or normative. I briefly argued that the three most commonly used semantic (realist-cognitivist, contextualist, and expressivist) and normative (consequentialism, contractualism, and virtue ethics) theories give us reason to use an everyday perspective, i.e., where we perceive people as responsi-

ble. This suggests that the answer to question 1, about whether we should give any weight to the skeptical intuitions elicited by the arguments from heteronomy and luck, is that we should not.

Of course, there is a lot to say about the explanation hypothesis and its implications. One objection we have often encountered is that there are cases where we judge people to be responsible even though nothing is unexpected about the motivational structure. If something is seen as a significant explanation of something else only insofar as it stands out from our expectations, we should not see these motivational structures as significant and thus should not, by the explanation hypothesis, see the agent as responsible for the action. Take the case I described in the summary of Paper 1, where a police officer fails to stop a robbery because she is more interested in the dog races than in the street life. In a corrupt society, for example, we would not expect her to help us, assuming we did not bribe her. Her inaction would be expected, but we still judge her to be responsible, in apparent conflict with the explanation hypothesis.

There are many ways of responding to these sorts of arguments, where one is to question whether we actually judge the agent to be responsible in these scenarios, and another is to argue that the explanation hypothesis actually gives the right result. Beginning with the first way of responding, we might have reason to hold police officers who are not doing their jobs without being bribed responsible, this to reform the system, making it less corrupt. However, if we do not think that the individual (in our case, the police officer) is doing anything wrong, and everyone else is doing it (taking bribes), then we might be inclined to think that perhaps the person is not fully responsible after all. Society, not the person, is to blame.

Furthermore, if it is the case that we actually think that, for instance, the police officer does something wrong, which probably most of us think, then we have *normative* expectations of that individual. Therefore, if she violates them, her motivational structure will become a significant explanation of there being something wrong with her behavior, and the case against the explanation hypothesis disappears. A related question, somewhat more difficult to handle, concerns cases where we perceive people to be responsible but do not violate our normative expectation. For example, a man takes his dog for a

walk because he is sensitive to its needs. Here it seems the person is responsible without any norm violation or anything else surprising going on. Even though it is an interesting question how the explanation hypothesis is to handle cases like this, I will not pursue it further here.

Another objection claims that the explanation hypothesis is somewhat “loose around the edges.” For example, many have asked about how we individuate motivational structures (for Reactive Response-ability). We readily admit that we have no precise account of this, partly for the reason that we think that people individuate motivational structures somewhat differently, and this is something that we continue to think about. However, a precise account is not required for saying something interesting about questions involving the explanation criteria (Explanation Responsibility), as we have done in Paper 1 and Paper 2.

A related objection, and the last one I will discuss here, is that it is impossible to test the explanation hypothesis empirically. This is related to the model’s vagueness and individual difference in the usage of explanatory frames. Of course, I disagree. The explanation hypothesis has been tested both directly and indirectly. The direct testing of the model has been done by Johansson (2010), which we accounted for in Paper 2. For instance, the subjects were asked to what extent people were responsible and at the same time to what extent they judged people’s motivational structure to serve as an explanation for the outcome at hand. The results were very much in line with what one would expect, given the explanation hypothesis. The explanation hypothesis has also received support indirectly by giving the right results when other theories have been tested and failed. (For more on this, see the next section). However, it is true that evidence in favor of the explanation hypothesis is a bit difficult to quantify, because it allows variations among the explanatory perspectives individuals use. This is difficult to amend in full. What we can do to add precision is to make predictions based on independently supported hypotheses about the relation between explanations and normative expectations. This is a work in progress.

The second question posed in section 3.2 was about whether we can give a unified account of the findings from experimental philosophy. We argued that the explanation hypothesis can provide such an account. The results in focus

were those about heteronomy and what I labeled in section 3.2.3 as moral luck (i.e., the Knobe effect). When experimental philosophers have examined lay people's intuitions regarding heteronomy, they have found that people's intuitions depend on how these cases are described. Cases that are described in some detail (concrete) and/or are of high affect, and/or are nonreductive (psychological) make people attribute responsibility to agents whose actions are ultimately determined by factors outside their control. Conversely, relevantly similar cases that are described abstractly without much or any detail and/or are of low affect and/or are reductive (neurological) make people judge those whose actions are ultimately determined by factors outside their control as excused. Furthermore, when experimental philosophers have examined lay people's intuitions regarding "moral luck," it has been found that "the folk" judge those who do things that have negative side effects as responsible, while those who do things with good side effects are not judged as responsible, although the other seemingly relevant circumstance is the same—the individual being judged does not care about the side effects, only the profits.

We argued in Paper 2 that the explanation hypothesis can give a unified account of these results. For example, describing a case concretely makes it much easier to see the motivational structure as part of a significant explanation than when it is described in an abstract way. Since we tend to perceive people as responsible when their motivational structures are judged as significant and as not responsible when the motivational structures are not, this distinction accounts for our divergent intuitions when subjected to abstract and concrete cases. The same is true for the other variations, such as those from high and low affect, and the reductive and nonreductive scenarios. We also argued that the explanation hypothesis can account for the Knobe effect. When someone brings about something bad without caring much about it, we would think of his motivational structure as explaining why the bad thing happened. On the contrary, when someone brings about something good without caring much about it, we would not think of her motivational structure as explaining the outcome. It seems like a fluke that the good thing happened, while it seems quite clear why the bad thing happened, and this explains the results in the Knobe-asymmetry.

There is a lot to say about our argument in Paper 2 as well. Besides the worries directly related to the psychological model discussed in the previous paragraphs, a flow of new empirical research on folk intuitions might shed new light and create potential problems for the explanation hypothesis. However, it seems that at least the most recent research actually does the reverse and strengthens the case for the explanation hypothesis, suggesting that the rival hypotheses are wrong. As an illustration, I will here address two of the most interesting and well-known alternative proposals built on recent empirical results: the *bypass hypothesis* (Nahmias and Murray 2012) and the *non-existence hypothesis* (Rose and Nichols *forthcoming*). I will discuss these in order.

Remember the bypass hypothesis from section 3.2.3, where Nahmias and Murray (2012) argue that the exculpating judgments from the Nichols and Knobe study (2007) could be explained with reference to the subject wrongly believing that determinism implied causal inefficacy of actions. If the bypass hypothesis is correct, we have a competing hypothesis to the explanation hypothesis, which at least has the advantage of being simpler and familiar (compatibilists have for centuries argued that incompatibilist intuitions come from conflating determinism and fatalism). The case for the bypass hypothesis might seem quite strong. For example, Murray and Nahmias (2012) have presented seemingly compelling evidence that people interpret the deterministic scenarios by Nichols and Knobe (2008) in a fatalistic way. Believing that people are excused in the deterministic scenarios strongly correlates with subjects believing that people's motivational structures in deterministic worlds are bypassed, such that their motivational structures have "no effect" on what they do (Murray and Nahmias 2012: 14). Furthermore, statistical analysis seemed to suggest that the belief in the bypassing explained these excusing intuitions (Nahmias and Murray 2012: 15–16). All this seemed to lend strong support to the bypass hypothesis.

However, more recent studies cast these results in doubt. According to the bypass hypothesis, people interpret some deterministic scenarios as implying bypassing, and since bypass judgments cause exculpatory judgments, they judge people not to be responsible in these cases. However, Björnsson (*forthcoming*) and Rose and Nichols (*forthcoming*) have found that responsibility

judgments (RESPONSIBILITY) drive the bypassing judgments (BYPASS) rather than the other way around. When we judge someone to be excused, we are, under certain conditions, also prone to judge the individual's motivational structure as bypassed. If this is correct, we cannot explain away the incompatibilist results after all, at least not through suggesting that people conflate determinism with bypassing.

For example, when using Nichols and Knobe's conditions with Universe A and Universe B (UNIVERSE) described in 3.2.2, Björnsson has found that "there was virtually no direct effect of UNIVERSE on BYPASS in a model treating RESPONSIBILITY as mediator" (Björnsson *forthcoming*: 19f). But when treating RESPONSIBILITY as a dependent variable and UNIVERSE as the independent variable, "77% of the total effect of independent on dependent variable was direct...suggesting that bypass interpretations play at most a partial role in explaining incompatibilist intuitions" (Björnsson *forthcoming*: 19). It seems, then, that responsibility judgments can explain bypass judgments, but not vice versa. Of course, we need more studies on the topic to be certain. Nonetheless, two studies refuting the bypass hypothesis, both involving a series of experiments, give us reason to believe that the bypass hypothesis is false.

Another rival of the explanation hypothesis is the *non-existence hypothesis* by David Rose and Shaun Nichols (*forthcoming*). This hypothesis was formulated after they found that responsibility judgments lead to bypassing judgments, not the other way around. The non-existence hypothesis states that when we judge determinism to be true, we are less likely to believe that any actual decisions are being made, which leads to both bypass ("agents' decisions have no effect on what they do") and exculpatory judgments. To test this, Rose and Nichols performed two experiments, finding evidence in favor of non-existence explaining bypass results (Rose and Nichols *forthcoming*: 13–18). For example, nearly half (47%) of the subjects believed that no decisions were made in the deterministic universe, and of course, if we cannot make any decisions, it seems likely that we are not responsible. However, a more recent study by Björnsson (*forthcoming*: 20) found the reverse direction of causation. The original idea was that the non-existence judgments caused the bypassing judgments. However, Björnsson's study suggests that bypassing judgments lead to non-existence judgments about decisions. I will not delve into details,

but Björnsson's (*forthcoming*) study casts doubts on the results of Rose and Nichols' research (*forthcoming*).

Björnsson (*forthcoming*) has also found support in favor of the explanation hypothesis. Björnsson hypothesizes that it is *not* that misreading of deterministic scenarios leads to bypass judgments, i.e., people think that the motivational structure has no effect on the outcome, which in turn leads to exculpatory judgments. Instead, people think that these motivational structures do not have any *independent effect* on the outcome, which in turn leads to exculpatory judgments. New data is in line with this hypothesis (Björnsson *forthcoming*). This result gives us further reason to believe in the explanation hypothesis, because if we take a motivational structure to have no independent effect, we are inclined to think that the structure does not provide significant explanations for outcomes. The end results then are twofold: we do not have any strong empirical evidence in favor of competing explanations, and the explanation hypothesis stands its ground well.

The third question was about whether policymakers and healthcare personnel should use an everyday or skeptical perspective. I argued that healthcare personnel should use the former while policymakers should use the latter. My argument appealed to the most commonly used normative considerations at the different levels—utilitarianism and egalitarianism at the policy level and beneficence, non-maleficence, and autonomy at the level of healthcare personnel. The difference in how they ought to perceive people chiefly concerns what they are able and likely to do at the different levels. Policymakers are at risk of issuing less effective policies if they perceive people as responsible, because they are less likely to set up primary preventive measures and thereby give people from vulnerable groups the help they need. Healthcare personnel dealing directly with their patients are more likely to violate patient autonomy if they do not perceive the latter as responsible, but are also less likely to give them the most effective treatment. Healthcare practitioners will be less inclined to advise patients on how to change their ways but are also less likely to include their patients in planning their treatment.

Of course, there is a lot to say about the argument in this paper as well. First, one could question whether the explanation hypothesis does any work here, and if it does, what this work is. I have discussed this question in Paper

3, but due to the scope of the paper, the issue might be easy to miss. One contribution of the explanation hypothesis is to explain why the discussion about responsibility and health in certain areas has the form it has and how we can resolve it. The World Health Organization, for instance, puts forward a skeptical concern similar to the arguments from heteronomy, focusing on prior causes and the surrounding society's effect on people, when arguing that people are not responsible for their ill health and therefore should not be held as such. Others in the debate focus instead on individual motivational structures and what people actually do when arguing that people are responsible for their ill health and should be held as such. The explanation hypothesis' relevance to this discussion is quite easy to recognize; it can explain the different sets of intuitions and suggest how this conversation can move forward.

The explanation hypothesis also forces us to reformulate the discussion about whether we should hold people responsible for their ill health. Traditionally, when asking whether we should hold people responsible, we examine different arguments for and against it, where one argument (that from responsibilitarianism) relates to whether people in fact are responsible. Given that the explanation hypothesis is true, we cannot approach the question about holding people responsible in this way. Instead, we first have to ask, at least if we follow the normative and not the metaethical track, whether we have a decisive normative reason to perceive people as responsible. Then if we do have this reason, we proceed to ask whether we should actually hold people responsible, bringing the notion that they are so and that we, because of it, have a *pro tanto* reason to hold them as such. Of course, other reasons might trump that of responsibility when discussing this issue, but the main point is that with the explanation hypothesis, we add one layer of questioning.

The fourth and last question was about whether we should prioritize bariatric surgery. I argued for a positive answer, after examining the vast amount of empirical work and combining it with the most common normative theories being used when setting priorities (the utilitarianism, the principle of greatest need, and the responsibilitarianism). Those who are generally eligible for bariatric surgery are severely obese, and they belong to those who have the greatest need. They suffer both from ill health and from an extremely low quality of life because of their weight. Furthermore, the obese also suffer from

significant health risks and the possibility of premature death in the future. Bariatric surgery is also the treatment of obesity that creates the most utility. First, this treatment is the only effective means by which we can help the obese to lose weight over the long term. Second, bariatric surgery is cost-effective, and the negative side effects are easily balanced by the positive direct effects. For example, some deaths, infections, etc., have occurred, but these are quite few and far between, and they also have to be weighed against what the majority gain. Lastly, people are not responsible for becoming or staying obese. There are many reasons for such a claim, but one has to do with non-culpable ignorance and the other with addiction and related problems. People generally do not know how hard it is to lose weight when they become obese or that it is easy to become addicted to sugar and fat. Furthermore, many of the obese are addicts or suffering from addiction-like states, making it almost impossible to control their eating in the long run, given their culture where the food to which people are addicted is marketed and available everywhere.

There is a lot to say about the main argument in this paper as well. However, the most general and interesting problem is to understand clearly the relevant aspects we want to measure and then make more refined statements about them. For example, I argue that the positive direct effects trump the negative direct impacts. Here, the positive direct effects come from losing weight and keeping it down. The negative direct impacts result from complications such as infections and bleeding. When I argue that the positive wins over the negative, I do so in part by citing medical expertise affirming it, but also on some sort of intuition based on empirical data. The extent to which people's health-related quality of life improves (which is significant) and the frequency at which the negative side effects occur (such as only a low percent suffer from greater problems) have been measured. From these findings, I have drawn the conclusion that undergoing bariatric surgery is generally worthwhile. In other words, I have not cited a measure that considers both positive and negative effects, puts them on the same scale, and finds the figure positive. What I have done most closely related to this is to cite findings that bariatric surgery is cost-effective by the National Health Service standards in terms of resource unit per quality-adjusted life years (QALY). However, QALY is known to have different sorts of problems, such as the uncertainty whether

it actually measures what we want (e.g., Brülde 2003). Something similar can be said about the other arguments as well, such as those from greatest needs and responsibility. The discussion about these things is a bit indeterminate. I will not try to resolve these difficulties here, but it is worth noting that they will have to be solved if we want to go further and issue more precise statements about what this “high priority” should amount to in terms of resource allocation.

4.2 IMPLICATIONS

There are both theoretical and practical implications of the research presented in this dissertation, especially implications of the explanation hypothesis. I will discuss some of these implications in order.

4.2.1 Theoretical implications

There are wide-ranging implications immediately acknowledged in understanding the explanation hypothesis. I will cover three implications here; the first two will be the arguments in favor of the different source views briefly mentioned in 3.1 and 3.2.2, those from manipulation and those against alternative possibilities (Frankfurt-style cases). Björnsson and I have discussed these two arguments in an early version of Paper 1, which is another reason why I take these up under *implications*. The third implication is that of collective responsibility, because it might be relevant to health. Björnsson has argued that the explanation hypothesis can account for our intuitions concerning collective responsibility, and I will give a short account of his line of reasoning here.

However, it should be noted that I will not supply any full-fledged descriptions of the arguments or exactly how the explanation hypothesis is supposed to account for these arguments. Instead, I will hint at what the explanation hypothesis can do in relation to the other arguments in the discussion about responsibility. Given that the explanation hypothesis can say something interesting about most of the relevant arguments, its validity is further strengthened. It also gives us reason to believe that the whole debate on the conditions for responsibility should be reframed. I will now discuss the argument from manipulation, Frankfurt-style cases, and collective responsibility, in that order.

From 3.1.1, recall Pereboom's four-case argument, with Professor Plum, who killed White. In the first scenario, Plum was directly controlled by the neuroscientists; in the second, he was indirectly controlled by neuroscientists; in the third, he was conditioned by society and culture; and in the fourth, his action was determined by factors outside his control. Intuitively and according to many compatibilists, Plum is not responsible in the first and perhaps not in the second case, whereas in the third and surely the fourth case, Plum is responsible. Pereboom's challenge to the compatibilist is to find a relevant difference between the first and the fourth cases so that the exculpation of responsibility from case 1 does not generalize to case 4. He argues that this cannot be done; hence, compatibilism is false.

A number of compatibilist strategies have been used to answer this objection. These can be said to be either hard-line or soft-line responses (McKenna 2008). The hard-liners argue that compatibilists should judge Plum to be responsible in case 1, but the soft-liners contend that we should try to find a relevant difference between the first and the fourth case, since Plum is not responsible in case 1. According to the hard-liner Michael McKenna (2008), if we "[f]ix attention on salient agential and moral properties" of Plum by examining the cases in reverse order and focusing on what case 1 has in common with case 2, this will make it seem considerably more plausible that Plum is indeed responsible in case 1.⁹³ Pereboom's response has been that we should focus equally on the different aspects of the case, both on the manipulation and the agential properties (Pereboom 2005, 2008). If we do that, he contends, we will judge Plum as not responsible.

Of course, the explanation hypothesis can easily explain both our initial intuitions when first confronting the four-case argument and the discussion between the hard-line compatibilists and Pereboom. In Pereboom's original description of the four-case argument, what straightforwardly explains Plum's actions in case 1 are the neuroscientists. They control Plum's every move and

⁹³ John Martin Fischer has argued for the hard-line position with similar arguments (2006: 230–34). He contends that we are responsible in all the cases but are not appropriate candidates for blame or praise in the first two cases. "After all, Plum is *not* a mere robot—he is not compelled or forced to act the way he does. He *does* exercise control, minimal as it may be" (Fischer 2006: 233).

make him commit a grossly immoral act, implying that their motivational structure emerges as a significant explanation for the murder.

Regarding the discussion between Pereboom and the hard-liners, we can see that they have already done most of the work for us. McKenna wants us to “focus” on Plum’s “agential properties,” i.e., his motivational structure, and Pereboom wants us to use a more general focus. According to the explanation hypothesis, we will perceive Plum as responsible if we indeed manage to focus on his agential properties. We will see an individual who performs a great moral transgression, which is something we (normatively) expect him not to do. On the other hand, if we apply a wider focus, Plum’s motivational structure will be thrust into the background and neurologists into the foreground, which again will make us perceive the neurologists as responsible. The explanation hypothesis also supplies us with a tool to help us decide which of these perspectives we should take. Should we focus on Plum’s motivational structure or that of the neurologists? It all boils down to what we should do according to the normative or metaethical theories that are correct.

The Frankfurt-style examples have also been briefly explained in 3.1.1. They are supposed to show that we do not have to be able to do or act otherwise or prevent an event from happening, etc., in order to be responsible for it. The classic version is when an individual is deliberating on doing something and does it because she wants it, thinks she one has a good reason for doing so, and so on. But unknown to the person, if she would have tried to do anything else, some nefarious neuroscientist would have made her do something else. This implies that one could not do otherwise. However, many people think that the person is responsible because he or she performed the action of his or her own accord, because she wanted to do so, thought she had reasons to do it, etc. That someone else was standing passively at the side and could have done something to change her ways of acting has no bearing on her responsibility for the action she actually performed.

Now, as we saw in 3.1.1, there has been a great debate about these Frankfurt-style examples. One approach has been to argue that we need to be careful about what we hold people responsible for, and when we examine the cases more thoroughly, we will realize that we actually do have to be able to act oth-

erwise to be responsible for the event at hand.⁹⁴ Begin with a case where all agree that we cannot be responsible for the event at hand, for instance, such as being struck with a terrible and debilitating genetic disease such as Huntington's. The development of Huntington's has nothing to do with how we act, where we live, etc. Instead, it is a purely genetic disease that we cannot do anything to avoid. The puzzle here becomes why we are not responsible for these unavoidable events, whereas we are so for other unavoidable incidents such as stealing cars, from the example in 3.1.1. In that case, whatever White tried to do, the nefarious neurologist Black would see to it that he ended up stealing the car. According to leeway incompatibilists' views, the puzzle can be solved through examining the cases more clearly. If we individuate things, we can see that we actually can avoid certain events, even though we cannot avoid others, and we are only responsible for the former, not the latter. White can avoid stealing the car in the relevant sense while we cannot avoid being struck with Huntington's, and this explains why we deem White to be responsible and those inflicted with Huntington's not to be responsible.

There are two types of events: consequence-particulars and consequence-universals. The difference between these two is that consequence-particulars can only be caused by one and the same thing, while consequence-universals can be caused by different things. We judge White to be responsible for the fact that the car was stolen because he could have avoided that "the car was stolen as a result of his own deliberation." After all, if he had tried to avoid stealing the car, the following would be true instead: "The car was stolen because of the neurologist's intervention" (cf. Fischer and Ravizza 1998: 96–98). Thus, it is simply false that the thief could not have avoided the consequence-particular for which we intuitively believe he is responsible. The consequence-universal, however, that "the car was stolen one way or another," was not something the thief could have prevented and not something for which he is responsible. This explains the puzzle about why some but not others seem to be responsible for inevitable states of affairs. No one is responsible for inevitable states of affairs; they only seem to be responsible for them.

⁹⁴ Inwagen (1978) was the first one to use this line of argument. See Fischer (2006) for a nice overview of that debate.

The other way to solve the puzzle of why some are responsible for inevitable consequences while others are not, according to the “conditionalists,” i.e., those who do not believe that we have to be able to act otherwise in an absolute sense in order to be responsible (see 3.1.1 for a more thorough description), is to argue that agents might be responsible for both consequence-particulars and consequence-universals when these are sensitive to our actions in the *actual sequence* that produced them (Fischer and Ravizza 1998: 98–115). The consequence-particulars and/or consequence-universals are sensitive to our actions in the right way when we can hold the “triggering event,” i.e., the event that for instance makes us do something we did not want to do, constant, and when doing this the consequence would not obtain if the agent did not act. Assuming that we can pick out these triggering events in some principled manner, we can see why agents are responsible for some “inevitable” consequence-universals/particulars but not others. For example, White is responsible for the consequence-universal that the car was stolen “one way or another,” because if White tried to do something else, and we held the neurologists’ device constant, the car would not have been stolen. The same is *not* true for people with Huntington’s disease. To be suffering from Huntington’s is not a consequence-universal sensitive to the afflicted individual’s actions. Those who suffer from it will do that, whatever they do. The explanation of the puzzle, then, is that people can only be responsible for events sensitive to people’s actions, etc., in the actual sequence that produced them. It should be noted, before we go on, that the different ways of approaching the subject, the absolutist and the conditionalist, have been thoroughly discussed since the late seventies.⁹⁵ Although great progress has been made, the best interpretation is that this discussion has now reached a stalemate.

Another well-known argument against the efficiency of Frankfurt-style cases has been launched by David Widerker (2000). Widerker recognizes the widespread intuition that the car thief is responsible for stealing the car, but he thinks that it disappears once we ask ourselves: *What should he have done instead?* When we contemplate that question and realize that there was nothing that the thief could have done instead, we are much less likely to feel that he is responsible for stealing the car. In response, McKenna (2005: 177) ar-

⁹⁵ For references, see footnote 94.

gues that although Widerker's question does affect intuitions, the original intuitions are reinstated *when we look at what he actually did*. We have thus reached what seems like a dialectical stalemate here as well.

It is easy to see how the explanation hypothesis can do some work in accounting for people's disparate intuitions and thereby take the discussion forward. Start with Widerker's argument that when we ask what the thief should have done instead, we would judge him not responsible. According to the explanation hypothesis, when we ask this question, the thief's motivational structure falls into the explanatory background. This is because the motivational structure is the same in all cases we reflect upon. We start comparing different scenarios, some where White chooses to act and the event occurs and others where White does not choose to act and the event occurs. For us to judge the motivational structure to be significant when asking contrastive questions, it cannot be the same in all the different scenarios under consideration. So when it is, as in these cases of overdetermination, we do not perceive the motivational structure as a significant explanation and hence judge the agent not to be responsible.

Furthermore, since the neurologist does something immoral here, forcing White to steal the car, his motivational structure stands out as significant because of its normative violation. Conversely, if we follow McKenna's lead, focusing on what actually happened, the motivational structure of the thief stands out again. We believe it is wrong to steal a car, so he violates our normative expectations; hence his motivational structure stands out as something that significantly explains the event at hand. Put differently, the contrastive questions are not asked, and hence we do not start looking for differences between different cases, which makes Whites motivational structure fall into the background.

Regarding the other discussion about absolutists' and conditionalists' views on whether we actually have to be able to avoid a certain event to be responsible for it, the accounting for this is also quite straightforward. We judge people as not responsible for illnesses not caused by their actions or omissions, such as purely genetic diseases, because the motivational structure is not, in any sense of the word, a factor among those explaining why the individual contracted the disease. The other diverging judgments of responsibility are, of

course, elicited by different sorts of framing effects. When we are confronted with the actual sequence and focus on what actually happened in the condition at hand, we perceive the person as responsible. The reason is that his or her motivational structure stands out as a significant explanation due to the fact that he or she, in our case, does something we expect him or her not to do. When we are taking the absolutist route, asking whether a person can be responsible for what happened, one way or the other, and we judge that the person cannot be responsible for universals, this is because our focus shifts when asking contrastive questions, and the motivational structure falls into the background due to the fact that it is the same in the different scenarios.

The last implication to be noted, collective responsibility, might be relevant to the discussion about public health and healthcare, because the cost for each obese individual, for instance, might be insignificant, but the cost for obesity as a whole is not. If there are collectives and they can be responsible and held accountable on such grounds, then we can legitimately hold the obese responsible because they inflict a huge cost on society, even though every obese individual per se does not. In terms of “restoration” (see 3.2.2), although no single individual creates scarcity, the collective does, and since the latter can be responsible, we have reason to hold it as such.

There are cases where we deem people jointly responsible for outcomes over which they as individuals have no control. Take the following case posed by Björnsson:

The Lake: Alice, Bill and Cecil each have a small boat in East Lake outside their town. One day last spring, each painted the boat and, unknown to the others, poured excess solvent into the lake. In the back of their heads, they all knew that this could affect the wildlife, but each of them decided that it would be a hassle to dispose of the solvent in a safe way and hoped that nothing bad would happen. However, as the solvent from all three diffused throughout the lake over the next few days, its concentration became high enough everywhere to prevent microorganisms in the lake from reproducing during the next few weeks, thus leaving higher organisms without food and effectively wiping out all fish in the lake. The concentration of solvent exceeded the threshold for the microorganisms by quite some margin: although the solvent from only one of the three would not have been enough to kill off the fish, the solvent from two would have. (Björnsson 2011: 181f)

In this case, most people would say that Alice, Bill, and Cecil are jointly responsible for the death of the fish. At the same time, it is quite clear that none of them performed an action necessarily or individually sufficient for the killing of the fish. The outcome would have happened even if one of the individuals had decided not to act the way he or she did, and the outcome would not have happened if only one had poured solvent into the lake. For example, if Alice had not poured her solvent into the lake, the effects would have been the same in the sense that the fish would still have died.

This is an intriguing case, then, which seems to hold that they as a group are responsible for the state of affairs, even though none of them is in control of the event at hand. There are two common explanations for arriving at these judgments: those from social ties/joint intentions (Held 1970, Kutz 2000, Miller 2006) and those from overlapping individual responsibility (Sverdlik 1987). First, we can see immediately that the lake case does not fit well with the idea that our judgment should be explained with reference to joint intentions or social ties. The individuals in the case do not know each other and are unaware that the others are doing the same thing they are doing; hence they do not share intentions or social ties. Second, there is also reason to believe that the notion of overlapping individual responsibility cannot do the proper work here either. This is because we would not perceive one of these individuals to be responsible if the others were replaced by nonagent causes. For example, if a large factory near the lake had discharged an extreme amount of solvent sufficient to poison the lake, then we would not consider Alice responsible for its poisoning, even though she poured solvent into it. This implies that we cannot understand joint responsibility as overlapping individual responsibility.

The explanation hypothesis can explain why we judge these individuals responsible for poisoning the lake. What seems to be going on here is that together they significantly explain the outcome of poisoning the lake. If they had been more careful, then the lake would not have been poisoned, and the fish would have survived. There is also no other subset giving a more satisfactory explanation of the event. If we were only to include Bill and Cecil when we try to explain why the fish died, we would seem to suggest that Alice had less part in the poisoning of the lake, which of course is wrong.

Going back to the obesity case, we could say that many people consider the obese responsible for the lack of resources in healthcare, for instance, because they think that the obese are a significant explanation for it. If they had cared more about their bodies, then we would not be in this position. Of course, this argument can be proven wrong, which I have partly done in Paper 4. Nonetheless, assuming it is correct (that the obese not caring enough is a significant explanation for the scarce resources in healthcare and perhaps also that holding them responsible would change their ways into eating healthier foods and exercising more), we might have reason to hold on to this perspective, where we perceive them to be jointly responsible. However, the overall point here is that the explanation hypothesis can account for cases where we judge people to have joint responsibility without them as individuals having control, and it also gives us a way to advance after these judgments have been made.

4.2.2 Policy implications

We have seen the wide range of theoretical implications from the ideas presented in this dissertation. There is also a broad spectrum of practical implications regarding the policy area. For example, we should not hold the obese responsible, but we should perceive policymakers as responsible. However, this discussion is highly diversified, so I will summarize all these different aspects relevant for policymaking and then state what my results imply for these various parts. In the end, this will amount to a checklist we can use when discussing questions about responsibility, accountability, and health.

The first point to consider is on what grounds we hold people responsible. According to some fundamental normative theories (such as responsibilitarianism), we have reason to hold people responsible for their ill health just because they are responsible for it. According to other normative theories (such as utilitarianism), however, we have no such reason to hold people responsible just because they are responsible for their ill health. Likewise, while according to some normative theories, we might have reason *not* to hold people responsible when they are excused, other normative theories do not put any weight

on people's innocence.⁹⁶ For example, if utilitarianism is true, and we would do the most good through holding people responsible for their ill health even though they were not responsible, because it would give them an incentive to adopt a healthier lifestyle in the future, then we have decisive reason to hold the "innocent" accountable. This line of reasoning amounts to the following:

The relation between being responsible and holding responsible:

- A) Being responsible \rightarrow *pro tanto* reason for holding responsible.
- B) Being responsible \rightarrow no *pro tanto* reason for holding responsible.
- C) Not being responsible \rightarrow *pro tanto* reason for not holding responsible.
- D) Not being responsible \rightarrow no *pro tanto* reason for holding responsible.

Given that the explanation hypothesis is true, this table gets even more complicated. Whether or not people are responsible hinges on whether we have a normative or metaethical reason to perceive these people as responsible. For example, we might have a decisive reason to judge people to be excused, and if this is the case, we also have reason to believe the arguments from moral luck or heteronomy to be true. This boils down to the following table:

The relation between being responsible and normative reasons:

- 1) Decisive normative reason to perceive as responsible \rightarrow Skeptical arguments rejected.
- 2) Decisive normative reason not to perceive as responsible \rightarrow Skeptical arguments accepted.

There is also a wide range of excuses, some of which we have discussed in this dissertation, most notably, those from nonculpable ignorance and addiction.

⁹⁶ A classic case against utilitarianism is about a sheriff who can do the most good through handing over an innocent man to a lynch mob (e.g., McCloskey 1963).

There are two views on nonculpable ignorance, the strong and the weak view, where the strong view implies that most people are excused, while the weak view does not. Regarding addiction, we have seen that under some conditions, it seems evident that it is responsibility undermining, while under other conditions, it is not. This leads to the following table:

Excusing conditions for responsibility:

- i) Addiction = Excuse.
- ii) Addiction = No excuse.

- iii) Weak view = most are probably culpably ignorant.
- iv) Strong view = most are probably nonculpably ignorant.

There is also a broad range of ways in which we can hold people responsible. We can make them pay in full or in part for their treatment, we can send reactive attitudes their way, we can deny them preventive work, and so on. Here are some of the most common ways of holding them responsible:

How we can hold people responsible:

- a) Financially liable, fully or partly.
 - *Ante facto*: mandatory insurance, taxes, etc.
 - *Ex post facto*: have them pay for the procedures, not paying for all the post-treatment, etc.

- b) Lower or no priority.
 - Examples: longer wait on transplant lists and in triage, fewer resources to lifestyle-based procedures, such as bariatric surgery.

- c) Reactive attitudes.
 - Examples: doctors scolding their patients who have taken unnecessary risks, society (government and nongovernment organizations) implementing blame campaigns aimed at groups involved in risky lifestyles, we as individuals blaming ourselves and others for risky behaviors.

- d) Prevention.

- Examples: put all resources on preventive efforts aimed at diseases of innocent children and others who are not responsible.

Here we should remember that the talk about “holding responsible” is meant to encompass both the positive and the negative senses of the word. Thus, all of the abovementioned things for which we can hold people responsible also come in positive versions. We should give people credit for choosing salad instead of fries at McDonald’s and give people who have taken responsibility for their health a higher priority in priority-setting contexts. The negative talk is much more common, however, which is why I use it here.

We can also be responsible for a wide range of different things, such as decisions, outcomes, and so on. This amounts to the following:

Things we can hold people responsible for:

- I) Positive and negative outcomes.
- II) Unnecessarily (directly or cumulatively) risky actions/omissions.
- III) Unnecessarily risky lifestyles.

We have also seen the different levels at which we can hold people responsible. Of course, this also follows from the discussion in the previous section of how we can hold people responsible, but it might still be good to state explicitly that there are different levels and roles; hence various rules of thumb, etc., might follow.

Levels at which we can hold people responsible:

- x) Macro: the highest political level where the government/parliament decides resource allocation to different sectors, the frameworks for the institutions, and so on.
- y) Meso: the middle level where the politicians/hospital directors, etc., decide to whom to distribute the resources, to give to more particular areas, and so on.

z) Micro: the lowest level, “the shop floor” where the healthcare personnel (i.e., doctors, nurses, administration) decide which patient should get a bed, which medication he or she should receive, and so on.

There is then a wide range of ways and things for which we can hold people responsible, on different levels, and so on. When discussing policy implications from what I have argued for in this dissertation, there are both new questions, such as 1 and 2, and new answers. I will now propose a checklist that policy-makers, including officers preparing decisions and decision makers themselves, can use when discussing responsibility, accountability, and health. This checklist will also be supplied with the answers given in this dissertation.

The first thing we have to ask in a situation where it might be appropriate to hold people responsible is whether we should perceive these people as excused, i.e., should we take the skeptical arguments to be viable? The first question in our checklist becomes:

1) Should we perceive X as excused?

We already have a partial answer to this question:

- On the macro level we should, as a rule of thumb, perceive people as excused.
- On the micro level we should, as a rule of thumb, perceive people not to be excused.⁹⁷

The second question in the checklist is whether people fulfill the other criteria for responsibility. For example, they might not be subjected to the arguments from heteronomy or luck, but they might still be addicts or nonculpably ignorant. The more general question then becomes:

2) Is X excused on other grounds than those accounted for by the explanation hypothesis?

⁹⁷ This rule, as we have seen, allows for exceptions within the frame of the explanation hypothesis, but these require the presence of special factors making excuse due to heteronomy or luck (small children, some psychiatric diseases, etc.).

Again, we have at least one answer to this question, illustrated by the case of obesity:

- People may be generally excused their ill health on the grounds of addiction and nonculpable ignorance (such as in the case of obesity).

The third question naturally concerns what, if anything, people are responsible for. This question is highly relevant because we might think it is fair to hold people responsible on some grounds but not others. For example, we might not think it to be fair to hold people responsible for individual injuries because we cannot know whether or not they are caused by factors outside the people's control. A broken leg might be the result of negligence but also arthritis. However, this does not imply that we cannot legitimately hold people responsible for anything. Presumably it is easier to find out whether people are indulging in risky lifestyles, and if we assume that people are responsible for choosing these, we can hold them accountable for that. The third question then becomes:

3) What is X responsible for?

We have seen above—in our general discussions, as well as those connecting to the specific case of obesity—that we could legitimately hold people responsible for outcomes, risks, and lifestyles, with the following implications:

- Outcome: We should be hesitant to hold people responsible for their ill health when addiction and/or nonculpable ignorance partly explain it, such as in the case of obesity.
- Risk: We should be hesitant to hold people responsible for cumulatively risky actions *ante facto* when these have resulted in such cases of ill health.
- Lifestyle: We should be hesitant to hold people responsible for risky lifestyles *ante facto* when these have resulted in such cases of ill health.

The fourth question is about holding people responsible. This is something that can be done or not, irrespective of what answers we have received to the questions above. Utilitarians might think it justified to hold the innocent responsible, and so on. This boils down to the fourth question:

4) Should we hold X responsible?

Of course, we have to be clear about what fundamental normative theories we are using when trying to answer this question. We also have one piece of the puzzle in answering question 3, which is:

- The obese should not be held responsible by being given lower priority in bariatric surgery.

Probably, the results will be more general than that, since we have seen that the obese have no alternative to bariatric surgery when trying to lose weight over the longer term. However, since I have not argued for that in this dissertation, I will not discuss it further. Likewise, the conclusion might extend also to other conditions and procedures which are relevantly similar, but again that is not implied directly by the analysis of the specific case of obesity and bariatric surgery.

The fifth question is about *how* we should hold people responsible. Even though it might be correct that people should not be perceived as excused on other grounds than those from heteronomy and luck, and we agree that we should hold them responsible, there might still be substantial disagreement on how and to what extent we should do so. Question five then becomes:

5) How should we hold X responsible in those cases when holding responsible is justified in general?

We have seen a number of ways of holding people responsible. Choosing which of these ways we should use comes down to what normative theory is correct. According to most theories, however, approaching people through holding them responsible for their ill health has to have some sort of positive effects, either for the individual approached or for others or society as a whole. This will exclude many ways of holding the obese and people in relevantly

similar sorts of ill health responsible, since it seems evident that such measures will not produce any positive effects. It does not help to reproach an obese individual for being such; he or she will not change, and it is plausible to assume the same to hold for other cases of ill health where addiction plays a role in a similar way. Furthermore, there are effective measures of preventing people from becoming obese, but blaming them, holding them financially liable, etc., do not seem to be among them (e.g., SBU 2002, 2004). The focus should instead be on population level and structural preventive measures such as promoting a healthy lifestyle in cities and villages by making nutritious and affordable foods available, providing safe access to sidewalks and designated lanes for walking and biking, offering balanced meals and physical exercise programs for school children, and so on.

However, in the future, we might encounter a legitimacy problem if we commit to giving the obese (or people suffering from some relevantly similar health problem) priority, because surgery and other possible solutions (i.e., new weight loss medicines) will incur considerable expense. This might drain resources from our healthcare institutions to the extent that others would deem it unfair for the obese to get such a large share of the budget. At least according to utilitarianism, we might reach a breaking point where we should let the obese or any others relevantly similar pay, at the minimum, a portion of their treatment expense. If people's trust in healthcare institutions plummets, they will not seek treatment in time and will not be as happy to contribute to government revenues with their taxes, which in turn may lead to negative consequences in terms of worse health among the population. Whether or not there actually is such a breaking point and, if so, where it is located is hard to say, but the possibility should be taken into account when discussing future prospects (as also should, of course, the possibility that there is no effect of this kind). Regardless of that, however, for the question of how we should hold the obese responsible today, the simple answer is that we should not.

4.3 FURTHER RESEARCH

The topic discussed in this dissertation calls for considerable further research. After all, almost every single one of the implications discussed above could (and should!) be elucidated more extensively to gain a better understanding of what the explanation hypothesis says regarding the Frankfurt-style cases, how the checklist for policy-related responsibility ascription and allocation should be developed to be of more substantial use in real policy work, etc. The main focus here, however, will be on the “reactive response-ability,” nonculpable ignorance, and the niceties of bariatric surgery.

A lot of work remains to be done to work out the details of the explanation hypothesis. If we review the comments in Paper 1, one task that seems most urgent is to clarify the reactive response-ability criteria of the explanation hypothesis in order to expand on what it actually entails. It should be noted that we might be able to use some examples, such as Fischer and Ravizza’s (1998) reason-responsive criterion. However, (as we noted in footnote 61), their criterion implies that people suffering from severe compulsive behavior or phobias might still be fully responsible. This position of Fischer and Ravizza is not only counterintuitive; it will probably be inconsistent with the implications of the explanation hypothesis. For example, if we take the metaethical route, examining the purpose of our practices of holding people responsible, it probably involves changing the motivational structure of the individuals we hold as such. However, if people suffer from severe compulsive disorders, it is highly unlikely that we will be able to change them in normal ways through, for instance, subjecting them to reactive attitudes. Therefore, the reason-responsive criterion of Fischer and Ravizza will not suffice, and we need to come up with another criterion, making the reactive reason-ability more clearly stated. When we get such a criterion in place, we not only have a more well-defined hypothesis about responsibility judgments, we also get a hypothesis that can be more easily tested. For example, we could perform experimental research to establish whether people agree with our general approach, if they judge people not to be responsible when the type of motivational structures is not sufficiently open to reform. If they do not share these intuitions, then it would count against the explanation hypothesis.

When we discussed culpable ignorance in section 3.3.2, it was quite obvious that the explanation hypothesis could provide clarification and explanation. The discussion is similar to the others mentioned about heteronomy, luck, etc. The intuitions we get in support of the different positions (the weak or the strong view) probably relate to how they aim our focus and hence affect our explanatory perspectives. For example, the strongest argument in favor of the weak view (which states that responsibility does not require an earlier point in time where the agent had an opportunity to become informed but because of *acrasia* reasons failed to do so) appeals to the case of a ruthless businessman committing bad acts, with a thorough description of his present state and his background that seemed nonexculpating. He is highly educated, he has met and meets people with views different from his, he has the overall capabilities to grasp moral reasons, and so on. This triggers the judgment that he is responsible according to the “explanation responsibility” criteria, because he does something immoral, i.e., he violates our normative expectations, and the explanatory perspective we are invited to take does not put focus on anything outside the agent explaining why he did what he did. Conversely, the strongest case for the opposite position makes it clear that the source of the faulty information ultimately lies outside the agent’s control. There is also an explicit question about what the agent in this situation should have done instead. When deliberating on what the agent should have done instead in the situation as it actually is, we are invited to take a first-person perspective. From this perspective, it seems that we could not have done anything, and hence the motivational structure falls into the explanatory background. Similarly, if it is made clear that the information ultimately was out of the hands of the agent, it seems evident that her motivational structure cannot explain why she did what he did.

It seems, then, that the explanation hypothesis in general, and the explanation responsibility criterion in particular, can say something interesting about the discussion on culpable ignorance. However, this discussion has to be expanded in much greater detail to make a real contribution. We have to examine the debate much more carefully to see what arguments are presented, whether the explanation hypothesis can account for these, and if so, how. The

answer regarding which perspective to take is, of course, also something that has to be worked out.

The last area worth further study is about obesity and what we should do about it in general and about bariatric surgery in particular. For example, it would be interesting to review the preventive works suggested and determine to what extent these can be implemented without sacrificing important moral values such as autonomy. It would also be good to know to what level we should prioritize the struggle against obesity. There are of course other worthwhile endeavors such as preventing the spread of infectious diseases and protecting youth from drug and alcohol addiction. It would also be interesting to examine further to what degree we should prioritize bariatric surgery. As for now, I have only mentioned that we should give it “high priority.” I have not discussed what it entails in terms of how to develop these practices in detail, for instance, whether healthcare professionals should actively reach out to the community to convince people who would not otherwise think about it to undergo surgery and thereby significantly increase the number of operations they perform. To do so, we would have to go through alternative treatments regarding other diseases, and so on, in greater detail. We probably also need to devise more definite measurements than the current ones in order to provide more noteworthy explanations about cases with only slight differences among the various treatments.

References

- Abraham K. (1993) Access to heart surgery for smokers. Wartime cigarette rations hooked a generation. *BMJ*, pp. 307:128e9.
- Adams, R. M. (1985) Involuntary Sins. *The Philosophical Review*, 94, pp. 3–31.
- Alexander, J. (2012) *Experimental Philosophy: An Introduction*. Polity Press.
- Anderson, E. S. (1999) What is the point of equality? *Ethics* 109 (2), pp. 287-337.
- Arneson, R.J. (1989) Equality and equal opportunity for welfare. *Philosophical Studies* 56, pp. 77–93.
- Arneson, R. J. (2011) Luck Egalitarianism –A Primer. In (ed.) Knight and Stemplowska *Responsibility and Distributive Justice*. Oxford U P, pp. 24-51.
- Atterbury, C. E. (1996) Anubis and the feather of truth: judging transplant candidates who engage in self-damaging behavior. *Journal of Clinical Ethics* 7, pp. 268e76.
- Aulisio M. P.; Arnold R. M. (1996). Exclusionary criteria and suicidal behavior: comment on “should a patient who attempted suicide receive a liver transplant”? *J Clin Ethics* 7, 277e83.
- Balaguer, M. (2010) *Free Will is an Open Scientific Problem*. The MIT Press.
- Bales, R. E. (1968) *Act- vs. Rule-Utilitarianism*. Thesis (Ph. D.)--Dept. of Philosophy, Stanford University.
- Beauchamp, T.L.; J.F. Childress (2008) *Principles of biomedical ethics*. Oxford U P.
- Benjamin M. (1997) Transplantation for alcoholic liver disease: the ethical issues. *Liver Transpl Surg* 3, 337e42.
- Berofsky, F. (1995) *Liberation from Self*. Cambridge U P.
- Björnsson, G. (2011) Joint Responsibility without Individual Control: Applying the Explanation Hypothesis. In *Compatibilist Responsibility: Beyond Free Will and Determinism*, eds. Jeroen van den Hoven, Ibo van de Poel and Nicole Vincent, Springer 2011, pp. 181–199.
- Björnsson, G. (forthcoming) Incompatibilism and “Bypassed” Agency. In Al Mele (ed.) *Surrounding Free Will*. Oxford U P
- Brülde, B. (1998) *Human good*. Acta Universitatis Gothoburgenes.
- Brülde B. (2003) *Teorier om livskvalitet* (Theories about quality of life). Student litteratur.
- Brülde, B.; Persson K. (2011) Hur skall de knappa resurserna fördelas? (How should

- the scarce resources be distributed?) In ed. Brülde, B. *Folkhälsoarbetets etik* (Public health ethics). Student litteratur.
- Buyx, A.M. (2008) Personal responsibility for health as a rationing criterion: Why we don't like it and why maybe we should? *Journal of Medical Ethics* 34, pp. 871–874.
- Cappelen, A. W.; Norheim, O. F. (2005) Responsibility in health care: a liberal egalitarian approach. *Journal of Medical Ethics* 31, pp. 476–480.
- Cappelen, H. (2012) *Philosophy without Intuitions*. Oxford U P.
- Casal, P. (2007) Why sufficiency is not enough? *Ethics* 55, pp. 296–326.
- Chisholm, R. (2003) 'Human Freedom and the Self,' *Free Will*, ed. M. Broukal and G. Watson, Oxford U. P., pp. 26–38.
- Clarke, R. (2003) *Libertarian Accounts of Free Will*. Oxford U P.
- Clarke, R. (2005). Agent Causation and the Problem of Luck. *Pacific Philosophical Quarterly* 86, pp. 408–421.
- Cohen, G.A. (1989) On the currency of egalitarian justice. *Ethics* 99: 906–944.
- G. A. Cohen (1985) Nozick on Appropriation. *New Left Review* 150, pp. 89–105.
- Cohen, G. A. (1995) *Self-Ownership, Freedom, and Equality*. Cambridge U P.
- Cole Wright, J.; Bengson, J. (2009) Asymmetries in judgment of responsibility and intentional action. *Mind and Language* 24 (1), pp. 24–50.
- Cova, F.; Bertoux, M.; Bourgeois-Gironde, S.; Dubois, B. (2012) Judgments About Moral Responsibility and Determinism in Patients with Behavioral Variant of Frontotemporal Dementia: Still Compatibilists! *Consciousness and Cognition* 21 (2), 851–64.
- Cuypres, S., E.; Haji, I. (2007) Authentic education and moral responsibility. *Journal of Applied Philosophy* 24 (1), pp. 78–94.
- Daniels, N. (2008) *Just health—meeting the needs fairly*. Cambridge U P.
- Daniels, N., Sabin, J. E. (2008) *Setting the limits fairly*. Oxford: Oxford U P.
- De Brigard, F (2010) If you like it, does it matter if it's real. *Philosophical Psychology* 23 (1), pp. 43–57.
- Deery, O.; Bedke, M.S. Nichols, S. (*forthcoming*) Phenomenal Abilities: Incompatibilism and the Experience of Agency In David Shoemaker (ed.), *Oxford Studies in Agency and Responsibility*. Oxford U P.
- Dennett, D. (1984) *Elbow Room: The Varieties of Free Will Worth Wanting*. Oxford U P.
- Dennett, D. (2003) *Freedom Evolves*. Penguin.
- Doris, J. M.; Knobe, J; Woolfolk, R. L. (2007) Variantism about responsibility. *Philosophical Perspectives* 21 (1), pp. 183–214.
- Driver, J. (2007) Attributions of Causation and Moral Responsibility," in Vol. II of *Moral Psychology*, ed. Walter Sinnott-Armstrong. MIT Press.

- Duus-Otterström, G. (2012) Weak and Strong Luck Egalitarianism. *Contemporary Political Theory* 11, 153-171.
- Dworkin, R. (1981) What is equality? Part 2: Equality of resources. *Philosophy and Public Affairs* 10, pp. 283–345.
- Dworkin, G. (1988) *The Theory and Practice of Autonomy*. Cambridge: Cambridge U P.
- Dworkin, R. (2002) *Sovereign Virtue –The Theory and Practice of Equality*. Harvard U P.
- Ekstrom, L. W. (2000) *Free Will: A Philosophical Study*. Westview Press.
- Elster, J. (ed.) (1999) *Addiction: Entries and Exits*. Russell Sage Publications.
- Engelhardt, T. H. (1996) *The Foundations of Bioethics*. Oxford U P.
- Faden, R. R; Beauchamp, T. L. (1986) *A History and Theory of Informed Consent*. Oxford U P.
- Feltz, A.; Cokely, E. T.; Nadelhoffer, T. (2009) Natural Compatibilism versus Natural Incompatibilism: Back to the Drawing Board. *Mind & Language* 24, pp. 1–23.
- Feltz, A.; Perez, A. (2012) Free Will, Causes, and Decisions Individual Differences in Written Reports. *Journal of Consciousness Studies*, 19 (9-10), pp. 166-189.
- Feltz, A. (2013) Pereboom and premises: Asking the right questions in the experimental philosophy of free will. *Consciousness and Cognition* 22 (1) pp. 53–63.
- Feinberg, J. (1970) *Doing & Deserving; Essays in the Theory of Responsibility*. Princeton: N.J., Princeton U P.
- Fischer, J. M. (1994) *The Metaphysics of Free Will*. Oxford: Blackwell.
- Fischer, J. M., Ravizza, M. (1998) *Responsibility and Control*. Cambridge U. P.
- Fischer, J. (2006) *My Way: Essays on Moral Responsibility*. Oxford U. P.
- Fischer, J.M.; Tognazzini, N.A. (2011) The Physiognomy of Responsibility. *Philosophy and Phenomenological Research* 82 (2), 381-417.
- FitzPatrick, W. J. (2007) Moral responsibility and normative ignorance: Answering a new skeptical challenge. *Ethics* 118, pp. 589–613.
- Fleck, L. M. (2009) *Just Caring: Health Care Rationing and Democratic Deliberation*. Oxford U P.
- Fojo, T., Grady, C. (2011) How much is life worth: cetuximab, non-small cell lung cancer, and the \$440 billion question. *J Natl Cancer Inst.* 5; 101 (15), pp. 1044-8. doi: 10.1093/jnci/djp177.
- Foddy, B., & Savulescu, J. (2010). A liberal account of addiction. *Philosophy, Psychiatry, and Psychology*, 17(1), 1–23.
- Frankfurt, H. (1969) Alternate Possibilities and Moral Responsibility. *The Journal of Philosophy* 66 (23), pp. 829–839.

- Frankfurt, H. (1971) Freedom of the Will and the Concept of a Person. *The Journal of Philosophy* 68 (1), pp. 5–20.
- Frankfurt, H. (1987) Equality as a moral ideal. *Ethics* 98 (1), pp. 21–43.
- Frankfurt, H. G. (1988) *The Importance of What We Care About: Philosophical Essays*. Cambridge U P.
- Gren, J. (2004) *Applying utilitarianism—The problem of practical action-guidance*. Acta Universitatis Gothoburgensis.
- Greene, J.; Cohen, J. (2004) For the Law, Neuroscience Changes Nothing and Everything. *Philosophical Transactions: Biological Sciences* 359, pp. 1775–85.
- Hare, R.M. (1981) *Moral thinking: Its levels, method, and point*. Oxford U P.
- Haji, I.; Cuyppers S. E. (2004) Moral responsibility and the problem of manipulation reconsidered. *International Journal of Philosophical Studies* 12 (4): pp. 439 – 464.
- Haji, I. (2009) *Incompatibilism's allure: Principal arguments for incompatibilism*. Peterborough: Broadview Press.
- Hart, H.L.A.. (1968) *Punishment and Responsibility—Essays in the philosophy of law*. Oxford U P
- Harris, J. (1995) Could we hold people responsible for their own adverse health? *Journal of contemporary health law and policy*, 147, pp. 147–153.
- Held, V. (1970) Can a Random Collection of Individuals be Morally Responsible?. *The Journal of Philosophy* 67, pp. 471–481.
- Higgs R. (1993) Should smokers be offered coronary bypass surgery? Human frailty should not be penalized. *BMJ* 306, pp. 1049e50.
- Hobart (1934) Free Will as Involving Determinism and Inconceivable Without It. *Mind* 43 (169), pp. 1–27.
- Holland, S. (2007) *Public health ethics*. Polity Press.
- Hooker (2000) *Ideal Code, Real World: A Rule-consequentialist Theory of Morality*. Clarendon Press.
- Johansson, E. (2010) *Testing the Explanation Hypothesis using experimental Methods*. Bachelors thesis, Linköping University. URI: urn:nbn:se:liu:diva-57308.
- Kagan, S. (1998) *Normative Ethics*. Westview Press.
- Kagan, S. (2012) *The Geometry of Desert*. Oxford U P.
- Kane, R. (1996) *The Significance of Free Will*. Oxford U. P.
- Kane, R. (1999) On Free Will, Responsibility and Indeterminism: Responses to Clarke, Haji, and Mele. *Philosophical Explorations* 2:2, pp. 105–121.
- Kane, R. (2005) *A Contemporary Introduction to Free Will*. Oxford U P.
- Kant, I. 1785 [2008] *The groundwork of the metaphysics of morals*. VA: Wilder.
- Kauppinen, A (2007) The rise and fall of experimental philosophy. *Philosophical Explorations* 10 (2), pp. 95 – 118.
- Knight, C. (2009) *Luck egalitarianism: Equality, responsibility, and justice*. Edinburgh

U P.

- Knobe, J. (2003) Intentional Action and Side Effects in Ordinary Language. *Analysis* 63, pp. 190–93.
- Knobe, J.; Doris, J. (2010) Responsibility. In John Doris & The Moral Psychology Research Group (eds.), *The Moral Psychology Handbook*. Oxford U P.
- Knowles, J. (1977) *Doing Better and Feeling Worse*. W.W. Norton.
- Korsgaard, C. 1996. Creating the kingdom of ends. Cambridge U P.
- Kutz, C. (2000) *Complicity*. Cambridge U P.
- Lewis, D. (1981) Are We Free to Break the Laws? *Theoria*, 47, pp. 113–21.
- Levy, N. (2003) Self-deception and responsibility for addiction. *Journal of Applied Philosophy* 20 (2), pp. 133–142.
- Levy, N. (2006) Autonomy and addiction. *Canadian Journal of Philosophy* 36 (3), pp. 427–447.
- Levy, N. (2009) Culpable ignorance and moral responsibility: A reply to FitzPatrick *Ethics* 119 (4), pp. 729–741.
- Levy, N. (2011) *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*. Oxford U P.
- Lie, L.M.; R.K. Sabik (2008) Priority setting in health care: Lessons from the experiences of eight countries. *International Journal for Equity in Health* 7 (4), 1–13.
- Martin, M. (2001) Responsibility for health and blaming victims. *Journal of Medical Humanities* 22 (2), 95–114.
- McKenna, M. (2004) Responsibility and Globally Manipulated Agents *Philosophical Topics* 32 (1/2), pp. 169–192.
- Mckenna, M. (2008) A hard-line reply to Pereboom's four-case manipulation argument. *Philosophy and Phenomenological Research* 77 (1), pp.142–159.
- McCloskey, H.J. (1972) A Non-Utilitarian Approach to Punishment.. In *Philosophical Perspectives on Punishment*. Ed. Gertrude Ezorsky. State University of New York Press, pp. 119–134.
- Mele, A. (1995) *Autonomous agents: From self control to autonomy*. Oxford U P.
- Mele, A. (2005) Libertarianism, Luck, and Control. *Pacific Philosophical Quarterly* 86, pp. 381–407.
- Mele, A. (2006) *Free Will and Luck*. Oxford U P.
- Mele, A. (2008) Manipulation, Compatibilism, and Moral Responsibility. *Journal of Ethics* 12, pp. 263–286.
- Mele, A. (2012) Another scientific threat to free will? *The Monist*, 95 (3) pp. 422–440.
- Miller, S. (2006) Collective Moral Responsibility: An Individualist Account. *Midwest Studies In Philosophy*, 30, pp. 176–193.

- Minkler M. (1999) Personal responsibility for health? A review of the arguments and the evidence at century's end. *Health education & behavior* 26, pp. 121-141.
- Moore, M. (1997) *Placing Blame: A Theory of the Criminal Law*. Clarendon Press.
- Moore, M. (2009) *Causation and Responsibility: An Essay in Law, Morals, and Metaphysics*. Oxford U P.
- Murray, D.; Nahmias, E. (2012) Explaining Away Incompatibilist Intuitions. *Philosophy and Phenomenological Research*, <http://dx.doi.org/10.1111/j.1933-1592.2012.00609.x>
- Nagel, T. (1979) Moral Luck in *Mortal Questions*. Cambridge U. P.
- Nahmias, E.; Morris, S.; Nadelhoffer, T.; Turner, J. (2004) The Phenomenology of Free Will. *Journal of Consciousness Studies* (11) 7–8: pp. 162–79.
- Nahmias, E.; Morris, S.; Nadelhoffer, T.; Turner, J. (2005) Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility. *Philosophical Psychology*, 18, pp. 561–84
- Nahmias, E.; Morris, S.; Nadelhoffer, T.; Turner, J. (2006) Is Incompatibilism Intuitive? *Philosophy and Phenomenological Research*, 73, pp. 28–53.
- Nahmias, E.; Coates, J.; Kvaran, T. (2007) Free will, Moral Responsibility, and Mechanism: Experiments on Folk Intuitions. *Midwest Studies in Philosophy*, 31, pp. 214–42.
- Nahmias, E., Murray, D. (2010) Experimental Philosophy On Free Will: An Error Theory For Incompatibilist Intuitions. In Jesus Aguilar, Andrei Buckareff and Keith Frankish (eds.) *New Waves in Philosophy of Action*, Palgrave Macmillan, pp. 189–216.
- Narveson, J. (2006) Justice in Health Care. *The Journal of Value Inquiry* 40, pp. 371–384.
- Nelkin, D. K. (2007) Do We Have a Coherent Set of Intuitions about Moral Responsibility? *Midwest Studies in Philosophy* 31, pp. 243–59.
- Nelkin, D. (2011) *Making Sense of Freedom and Responsibility*. Oxford U P
- Nelkin, D. K. (2013) Moral Luck, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2013/entries/moral-luck/>.
- Nichols, Shaun (2004) The Folk Psychology of Free Will: Fits and Starts. *Mind & Language* 19, pp. 473–503.
- Nichols, S; Knobe, J. (2007) Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions. *Nous* 41 (4), pp. 663–685.
- Nord, E. (1999) *Cost-Value Analysis in Health Care: Making Sense out of QALYS*. Cambridge U P.
- Nozick, R. (1974) *Anarchy, State and Utopia*. Basic Books.

- Nussbaum, A. (1992) Human Functioning and Social Justice: In Defense of Aristotelian Essentialism. *Political Theory* 20 (2), pp. 202-246.
- O'Connor, T. (1995) Agent causation. In *Agents, Causes, and Events: Essays on Indeterminism and Free Will*. Oxford U P.
- O'Connor, T. (2000) *Persons and causes: The metaphysics of free will*. Oxford U P.
- Olsaretti, S. (ed.) (2003) *Desert and Justice*. Oxford U P
- Olsen, D.P. (1997) When the patient causes the problem: the effect of patient responsibility on the nurse-patient relationship. *J Adv Nurs* 26, pp. 515e22.
- Parfit, D. (1984) *Reasons and persons*. Oxford U P.
- Parfit, D. (1991) *Equality or priority?*, 1–42. The Lindley Lecture: University of Kansas Press.
- Parfit, D. (2012) *On What Matters*. Oxford: Oxford U P.
- Pereboom, D. (1995) Determinism al dente. *Noûs* 29 (1), pp. 21-45.
- Pereboom, D. (2001) *Living without Free Will*. Cambridge U. P.
- Pereboom, D. (2005) Defending Hard Incompatibilism. *Midwest Studies* 29, pp. 228-47.
- Pereboom, D. (2007) A Hard-Line Reply to the Multiple-Case Manipulation Argument. *Philosophy and Phenomenological Research* 77, pp. 160-70.
- Pereboom (forthcoming) *Free Will, Agency, and Meaning in Life*. Oxford: Oxford U P.
- Persson, I. (2005) *The Retreat of Reason: A Dilemma in the Philosophy of Life*. Oxford: Oxford U P.
- Persson, K. (2012) Why Bariatric Surgery Should be Given High Priority: An Argument from Law and Morality. *Health Care Analysis*. DOI 10.1007/s10728-012-0216-1
- Pettit, D.; Knobe, J. (2009). The Pervasive Impact of Moral Judgment. *Mind & Language* 24, pp. 586–604.
- Rakowski, E. (1991) *Equal justice*. Clarendon.
- Rawls, J. (1971) *A theory of justice*. Harvard U P.
- Rose, D.; Nichols, S. (forthcoming) The Lesson of Bypassing. *Review of Philosophy and Psychology*.
- Rosen, G. (2004) Skepticism about moral responsibility. *Philosophical Perspectives* 18 (1), pp. 295–313.
- Roskies, A. L.; Nichols, S. (2008) Bringing Responsibility Down to Earth. *Journal of Philosophy*, 105, pp. 371–88.
- Sarkissian, H.; Chatterjee, A.; De Brigard, F.; Knobe, J.; Nichols, S., Sirker, S. (2010) Is Belief in Free Will a Cultural Universal? *Mind & Language* 25, pp. 346–58.
- SBU. (2002) *Obesity—problems and solutions* (Fetma –problem och åtgärder). Available at: <http://www.sbu.se/sv/Publicerat/Gul/Fetma—problem-och-atgarder/>.

- SBU (2004) *Preventive Measures Against Obesity* (Förebyggande åtgärder mot fetma). Available at: <http://www.sbu.se/sv/Publicerat/Gul/Forebyggande-atgarder-mot-fetma/>
- Schmidt, H. (2009) Personal responsibility in the NHS Constitution and the social determinants of health approach: competitive or complementary?. *Health Economics, Policy and Law*, 3 (2), pp. 129-138.
- Schmidt H., Gerber A., Stock S. (2009) What can we learn from German health incentive schemes? *British medical journal* 339, pp. 725-728.
- Schwitzgebel, E. Cushman, F. (2012) Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers. *Mind and Language* 27 (2), pp. 135-153
- Segall, S. 2009. *Health, luck and justice*. Princeton: Princeton U P.
- Sen, A. (1992) *Inequality Reexamined*. Oxford U P.
- Sharkey, K., and L. Gillam. 2010. Should patients with self-inflicted illness receive lower priority in access to healthcare resources? Mapping out the debate. *Journal of Medical Ethics* 36: 661–665.
- Shepherd, J. (2012) Free will and consciousness: Experimental studies. *Consciousness and Cognition*, 21, pp. 915–927
- Sher G. (2009) *Who Knew? Responsibility without awareness*. Oxford: Oxford U P.
- Sidgwick, H. (1907) *The methods of ethics*. Macmillan.
- Singer, P. (2009) Why We Must Ration Health Care. *New York Times*. Available at: http://www.nytimes.com/2009/07/19/magazine/19healthcare-t.html?pagewanted=all&_r=0. Accessed December 19th 2013.
- Smart, B. (1994) Fault and the allocation of spare organs. *Journal of Medical Ethics* 20 (1), pp. 26-30.
- Smith, H. (1983) Culpable ignorance. *Philosophical Review* 92 (4): pp. 543-571.
- Smith, M. (2003) Rational Capacities, or: How to Distinguish Recklessness, Weakness, and Compulsion. In Stroud and Tappolet, eds., pp. 17–38.
- Smilansky, S. (2000) *Free Will and Illusion*. Oxford: Oxford U. P.
- SOU. 1995:5. *Difficult choices in health care* [in Swedish] (Vårdens svåra val). <http://www.regeringen.se/sb/d/108/a/25124>. Accessed 13 Feb 2012.
- Sripada, C. S. (2012). What Makes a Manipulated Agent Unfree? *Philosophy and Phenomenological Research* 85 (3), pp. 563-593.
- Strawson, Peter F (1982) [1962] Freedom and Resentment. In ed. Gary Watson *Free Will*. Oxford University Press, pp. 59–80.
- Strawson, G. (1986) *Freedom and Belief*. Oxford: Oxford U. P.
- Strawson, G. (2001) The bounds of freedom' in *The Oxford Handbook on Free Will*, ed. R. Kane (Oxford University Press), pp 441–60.
- Sverdlik, S. (1987) Collective Responsibility. *Philosophical Studies* 51, pp. 61-76.

- Temkin, L.S. (1993) *Inequality*. Oxford: Oxford U P.
- Temkin, L. (2011) Justice, Equality, Fairness, Desert, Rights, Free Will, Responsibility, and Luck. In *Distributive Justice and Responsibility*, edited by Carl Knight and Zofia Stemplowska, pp. 51-76, Oxford University Press.
- Thomson, J. J. (1993) Morality and Bad Luck. In *Moral Luck*, D. Statman (ed.), Albany: State University of New York Press.
- Tonti-Filippini N. A. (2004) Smoking cessation and elective surgery: the cleanest cut a response. *Med J*, 181 pp. 283e4.
- Tännsjö, T. (1998) *Hedonistic Utilitarianism*. Edinburgh: Edinburgh U P.
- Tännsjö, T. (2006) Who Should Bear the Costs of IVF: In Search of a Just Solution. *Ethics, Science, and Moral Philosophy of Assisted Human Reproduction*, 13 Supplement RBM Online.
- Übel, P. A.; Jepson, C.; Baron, J., et al. (2001) Allocation of transplantable organs: do people want to punish patients for causing their illness? *Liver Transplantation* 7, pp. 600e7.
- Vallentyne, P.; Tungodden, B. (2006) Who are the least advantaged?. In *Egalitarianism: New Essays on the Nature and Value of Equality*, ed. Nils Holtug and Kasper Lippert-Rasmussen. Oxford University Press, pp. 174-195.
- Vallentyne, P. (2007) Brute Luck Equality and Desert. In Olsaretti ed. *Desert and Justice*. Oxford: Oxford U P.
- Vallentyne, P. (2008) Brute Luck and Responsibility. *Politics, Philosophy & Economics* 7, pp. 57-80.
- van Inwagen, P. (1983) *An Essay on Free Will*. Clarendon Press.
- Vargas, M. (2005) The Trouble with Tracing. *Midwest Studies in Philosophy* 29 (1), pp. 460-475.
- Wallace, R. J. (1999) Addiction as defect of the will: Some philosophical reflections. *Law and Philosophy* 18 (6), pp. 621–654.
- Walker, M. U. (1991) Moral Luck and the Virtues of Impure Agency. *Metaphilosophy*, 22, pp. 14–27.
- Walker, T. (2010) Who do we treat first when the resources are scarce? *Journal of Applied Philosophy* 27(2), 200–211.
- Walker, M. J. (2010) Addiction and Self-Deception: A Method for Self-Control? *Journal of Applied Philosophy* 27 (3), 305-319.
- Wallace, R. J. (1994) *Responsibility and the Moral Sentiments*. Harvard U P.
- Wallace, R. J. (1999) Addiction as defect of the will: Some philosophical reflections. *Law and Philosophy* 18, pp. 621–654.
- Watson, G. (1996) Two Faces of Responsibility. *Philosophical Topics*, 24 (2), pp. 227–48.

- Vargas, M. (2005) The trouble with tracing. *Midwest Studies in Philosophy* 29 (1), 269–291.
- Watson, G. (1999) Disordered Appetites: Addiction, Compulsion and Dependence. In Jon Elster (ed.), *Addiction: Entries and Exits*. Russell Sage Publications.
- Watson, G. (2004) *Agency and Answerability: Selected Essays*. Oxford U P.
- World Health Organization. (2008) *Closing the gap in a generation: Health equity through action on the social determinants of health*. http://www.who.int/social_determinants/en/. Accessed 13 Feb 2012.
- Widerker, D., McKenna, M. (eds.), (2003). *Moral Responsibility and Alternative Possibilities*, Aldershot, UK: Ashgate Press.
- Widerker, D. (2000) Frankfurt's Attack on the Principle of Alternative Possibilities: A Further Look. *Philosophical Perspectives*, pp. 181-201.
- Wikler, D. (2002) Personal and social responsibility for health. *Ethics & International Affairs* 16 (2), 47–55.
- Williams, B. (1981) *Moral Luck*. Cambridge U P.
- Williams, A. (1997) Intergenerational Equity: An Exploration of the 'Fair Innings' Argument. *Health Economics* 6, pp. 117–132.
- Wolf, S. (2001) The Moral of Moral Luck. *Philosophic Exchange*, 31, pp. 4–19.
- Wolff, J. (1991) *Robert Nozick: Property, Justice, and the Minimal State*. Stanford U P.
- Wolff, J. (2011) *Ethics and Public Policy: A Philosophical Inquiry*. Oxon: Routledge.
- Woolfolk, R. L; Doris, J. M.; Darley, J. M. (2006) Identification, Situational Constraint, and Social Cognition: Studies in the Attribution of Moral Responsibility. *Cognition* 100, pp. 281–301.
- Young, L.; Nichols, S.; Saxe, R. (2010) Investigating the Neural and Cognitive Basis of Moral Luck: It's Not What You Do but What You Know. *Review of Philosophy and Psychology* 1 (3), pp. 333-349.
- Zimmerman, M. J. (1997) Moral Responsibility and Ignorance. *Ethics* 107, pp. 410-426.
- Zimmerman, M. J. (2002) Taking Luck Seriously. *The Journal of Philosophy* 99, pp. 553–576.

Papers I–IV

Paper 1: The explanatory component
of moral responsibility

Björnsson, Gunnar;

Persson, Karl

Noûs 2012

46 (2) pp. 326—354.



I

Paper 2: A Unified Empirical Account
of Responsibility Judgments

Björnsson, Gunnar;

Persson, Karl

Philosophy and Phenomenological Research

2013, 87 (3): 611—639.



II

Paper 3: The Right Perspective on
Responsibility for Ill Health

Persson, Karl

Medicine, Health Care, and Philosophy

2012, 16 (3): 429—421.



III

Paper 4: Why Bariatric Surgery Should be
Given High Priority: An Argument
from Law and Morality.

Persson, Karl

Health care analysis

2012, DOI: 10.1007/s10728-012-0216-1.

IV

Svensk sammanfattning

Bör vi hålla människor ansvarig för deras dåliga hälsa? Denna fråga har diskuterats allt mer intensivt under senare år bland annat som en konsekvens av ökad resursknapphet och ökade kostnader för olika livsstilsrelaterade sjukdomar. Frågan är emellertid om vi är ansvariga för vår dåliga hälsa ens när det är fastställt att det faktiskt är livsstilen som har orsakat sjukdomen samt om vi i så fall bör hålla människor ansvariga för det. Om detta handlar den här avhandlingen.

Mer precist behandlar avhandlingen fyra olika frågeställningar i fyra olika artiklar. Den första artikeln (Björnsson och Persson 2012) behandlar de skeptiska argument som vanligtvis riktas mot vår förmåga att kunna vara ansvariga för något över huvud taget. Här diskuterar vi de vanliga heteronomi- och turargumenten och kommer fram till att de inte har något bevisvärde som argument mot vår förmåga att kunna vara ansvariga. Vi kommer fram till att de generellt inte har detta genom att ställa upp en psykologisk hypotes, *Förklaringshypotesen*, som klargör varför vi faller de omdömen vi gör när vi konfronteras med skeptiska argument.

Enligt förklaringshypotesen får de skeptiska argumenten oss att inta förklaringsperspektiv (förklaringsramar) där agenternas motivationella strukturer (motivation) inte betraktas som signifikanta förklaringar till de utfall vi söker förklaringar till. När vi intar dessa perspektiv finns inget hos agenternas motivation som vi upplever som överraskande eller som på annat sätt tar tag i vår uppmärksamhet. Samtidigt finns det inget som säger att dessa förklaringsperspektiv är något vi bör använda framför andra mer vardagliga perspektiv där agenternas motivation framstår som signifikanta förklaringar och där vi därför bedömer agenterna som ansvariga. Vi avslutar artikeln med att argumentera för att de vanligaste semantiska och normativa teorierna ger oss skäl att tro att vi generellt bör ta ett vardagligt perspektiv på människors handlande.

Den andra artikeln bygger delvis vidare på den första och undersöker huruvida det finns en enhetlig förklaring till de tillsynes disparata resultat rörande icke filosofers ansvarsomdömen som experimentell filosofi har visat (Björnsson och Persson 2013). Experimentell filosofi är en ny gren inom filo-

sofi där man främst undersöker om icke-filosofier reagerar på samma sätt som filosofer när de konfronteras med olika argument och tankeexperiment samt om det finns några underliggande psykologiska mekanismer som kan förklara dessa reaktioner. Artikeln visar att förklaringshypotesen kan ger klarhet i varför människor har dessa tillsynes disparata reaktioner.

Vi diskuterar två huvudsakliga fall som förklaringshypotesen kan klargöra: fall av heteronomi och fall av asymmetri. I heteronomifallen finns det en rad vinjetter som är lika varandra i relevanta avseenden men som människor reagerar olika på med avseende på ansvar. I samtliga vinjetter beskrivs människors handlande som fullständigt orsakat av faktorer utanför deras kontroll. Vissa beskrivningar, exempelvis de som är konkreta snarare än abstrakta, skrivna med hög affekt snarare än låg eller uttryckta i psykologiska snarare än neurologiska termer, gör dock att människor i betydligt högre utsträckning betraktar personerna i vinjetterna som ansvariga.

Skälet till dessa disparata reaktioner är enligt förklaringshypotesen att vi i de första fallen intar ett förklaringsperspektiv där agenternas motivation framstår som en signifikant förklaring till utfallet och att vi i de senare fallen intar ett förklaringsperspektiv där agenternas motivation inte framstår som en signifikant förklaring. När man till exempel ger abstrakta beskrivningar av fall där människor handlande ytterst är orsakat av faktorer utom deras kontroll intar vi ofta ett perspektiv där motivationen inte bedöms som en signifikant förklaring vilket medför att vi upplever att agenten inte är ansvarig. När man istället ger en konkret beskrivning av fall där människor handlande och så vidare beskrivs i större detalj, intar vi oftare ett mer vardagligt perspektiv där vi bedömer motivationen som en signifikant förklaring, vilket medför att vi upplever agenten som ansvarig.

I asymmetrifallen är frågan istället varför vi bedömer personer som medvetet gör något som har dåliga effekter som ansvariga, samtidigt som vi bedömer personer som medvetet gör något som har goda effekter som ej ansvariga. Det paradigmatiska fallet gäller en vd som får veta att ett projekt kommer att inverka negativt (positivt) på miljön. Hen förklarar att: ”jag bryr mig inte om miljön, jag bryr mig bara om att tjäna pengar”, och beslutar sig sedan för att sätta igång projektet. Människor som läser detta tenderar att bedöma vd:n som ansvarig för det dåliga utfallet, det vill säga när projektet skadar miljön,

men inte i det goda utfallet, det vill säga när projektet gynnar miljön, trots att inget annat skiljer fallen åt.

Enligt förklaringshypotesen är denna asymmetri förväntad. Normalt betraktar vi människors brist på omsorg av värdefulla saker som en förklaring till varför dessa förstörs. Vidare har vi normativa förväntningar på människor att försöka undvika att förstöra (den värdefulla) miljön. Detta leder till att vi lätt intar ett förklaringsperspektiv där vd:ns motivation framstår som en signifikant förklaring när dennes brist på omsorg leder till att miljön förstörs. Omvänt tänker vi oss normalt sett inte att en brist på omsorg av värdefulla saker gynnar dem. Vidare har vi förväntningar på människor att försöka gynna (den värdefulla) miljön. Detta gör att vi i fallet där miljön gynnas lätt intar ett perspektiv där vd:ns motivation inte betraktas som en signifikant förklaring och där vi därmed inte upplever hen som ansvarig.

I den tredje artikeln tar jag steget över till det mer konkreta fallet med människors ansvar för sin dåliga hälsa (Persson 2012). Jag argumenterar här för att vi antingen kan inta ett abstrakt perspektiv och betrakta människor som huvudsakligen ”beroende variabler” och därmed ej ansvariga, eller inta ett vardagligt perspektiv där vi betraktar dem som huvudsakligen ”oberoende variabler” och därmed åtminstone ej uteslutna från ansvar. (Det finns andra orsaker än de skeptiska argumenten som exempelvis de från oklandervärd okunskap och bristande impuls kontroll, vilket gör att de kan vara ursäktade även om vi använder vardagliga perspektiv). Jag finner att man bör inta olika perspektiv på olika nivåer, ett på policynivån och en annat på ”sjukhusgolvet”. Detta med utgångspunkt i de vanligaste normativa överväganden man använder sig av på de olika nivåerna.

På policynivån bör man ta ett abstrakt perspektiv och betrakta människor som ej ansvariga för sin dåliga hälsa. Det huvudsakliga argumentet är att det finns goda skäl att tro att betrakta människor på detta sätt kommer medföra bättre konsekvenser och ett mer jämlika utfall. De som betraktar människor som ej ansvariga för sin dåliga hälsa är exempelvis enligt empiriska studier mer benägna att sätta in preventiva åtgärder, vilka har visat sig vara mest effektiva för att förbättra folkhälsan. På samma sätt är de som betraktar människor som ej ansvariga för sin dåliga hälsa mer benägna att hjälpa de som har det sämst innan det gått riktigt illa, vilket leder till ökad jämlikhet. På sjukhusgolvet bör

man däremot betrakta människor som huvudsakligen oberoende variabler, och därmed vara mer benägna att tro att människor är ansvariga för sin dåliga hälsa. Detta eftersom risken minskar att man kränker människors autonomi men också för att man antagligen i högre grad kommer att involvera dem i de beslut som rör deras behandling. Förutom att det är effektivt ur ett kortare behandlingsperspektiv, ökar det sannolikt ansträngningarna att förändra deras beteende på längre sikt vilket också är gynnsamt för patienterna.

I den fjärde artikeln behandlar jag frågan om överviktskirurgi (bariatrisk kirurgi) bör prioriteras i förhållande till andra behandlingar. Jag argumenterar för att denna behandlingsform mot fetma bör få hög prioritet baserat på tre skäl.

För det första har gravt överviktiga bland de största behoven av sjukvårdens patienter. Personer som lider av fetma har i jämförelse med övriga populationen låg livskvalitet, dålig hälsa och stor risk för ännu sämre framtida livskvalitet och för tidig död. För det andra är överviktskirurgi den enda metod vi har evidens för att den fungerar på längre sikt. Konventionella metoder, som exempelvis kostrådgivning, har visat sig vara effektiva på kort sikt, det vill säga under ett till två år. Efter denna period går patienten i allmänhet upp till sin ursprungliga vikt. Kirurgi har däremot visat sig leda till varaktiga resultat. Överviktskirurgi är dessutom kostnadseffektiv enligt de vanligaste måtten, och har en låg grad av negativa sidoeffekter som dödsfall, blödningar, eller allvarliga infektioner. För det tredje menar jag att de som lider av fetma inte är ansvariga för sin grava övervikt. På populationsnivå misslyckas människor i allmänhet med att gå ner i vikt även om de anstränger sig. En förklaring till varför de misslyckas är att de ofta har ett missbruksförhållande till fet och sockerrik mat. Detta missbruk är hart när omöjligt att komma ur eftersom det omgivande samhället är fyllt med budskap som triggar igång begär och det finns social press att äta viss sorts onyttig mat. Vidare är inte kunskapen spridd varken om hur svårt det är att gå ner i vikt när övervikt väl nåtts eller om hur lätt kan komma att drabbas av denna typ av missbruk. Detta gör att de som lider av fetma idag inte kan sägas ha vetat vad de gav sig in i när de gick upp i vikt, vilket implicerar att de ej heller är ansvariga för detta.

Avslutningsvis ger avhandlingen en ny ingång i hur vi bör diskutera ansvarsfrågor generellt men den säger också något om det konkreta fallet om

ansvar och hälsa. Detta kan förhoppningsvis föra den viktiga och alltmer omdiskuterade frågan om hur vi bör förhålla oss till människor som lider av dålig hälsa till följd av sin livsstil framåt.