



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

“Internal Reasons and Rational Persuasion”

verfasst von / submitted by

Mag. Christian Demmelbauer BA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Arts (MA)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt /
Degree programme code as it appears on
the student record sheet:

UA 066 941

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Philosophie UG 2002

Betreut von / Supervisor:

Univ.-Prof. Dr. Hans Bernhard Schmid

Acknowledgments

Many people have contributed to making this text what it is and I feel deeply indebted to all of them. Though writing a text can be a rather solitary experience, it is an activity one cannot do without the support and influence of others. I was very lucky to be surrounded by many people who generously shared their thoughts and perspectives with me.

Sarah Espinosa, Gareth Pearce, Nikolaus Peschl and Marlene Valek provided extensive and insightful comments on earlier drafts in seminars held by Prof. Schmid at the University of Vienna in the winter of 2020 and in the summer of 2021. I am grateful for their patience with the obvious flaws these early drafts had and their probing questions and remarks which often guided me in the further development of my thoughts. I am equally thankful to the other participants of these seminars for challenging discussions of these drafts.

I am thankful to Martin Niederl and Hannah Romano for reading (parts of) drafts of this thesis, providing thoughtful comments and drawing my attention to mistakes I had made. Especially Martin Niederl undertook a careful reading of a rather late draft and generously offered many wise questions and comments, exposing many weaknesses and incomprehensibilities in my arguments. If one reads the footnotes carefully, one will find many traces of his reading in this text.

I am also deeply indebted to my supervisor Prof. Hans Bernhard Schmid. Especially in the beginning, he has closely accompanied my work, doing much to make my rather confused thoughts clear. His critical reading of later drafts allowed me to figure out much more precisely what the essence of my argument was and how I could put things more clearly.

It is not only academic support that is essential in completing a rather large text like this. My friends and my parents have done much to provide a loving and caring environment for me while at the same time challenging me in numerous discussions and debates. Especially my wonderful mother provided me both with an environment in which I could let go of the work and motivation to take it up again when it seemed to be drowned out by other commitments.

Lastly, and most importantly, I want to thank my partner Christina. This has been a longer process than it could have been, and she certainly had to suffer most from my moods during writing. Not once, however, did her patience with me come to an end. She has also read earlier drafts, exposed my mistakes, and discussed ideas contained in this thesis with me, showing me, time and again, that wisdom does not need fancy philosophical vocabulary to express itself.

Contents

I.	Introduction.....	1
II.	Internalism and Rational Persuasion	4
1.	Giving Reasons	4
a.	Force and persuasion.....	5
b.	Two types of reason-giving and the toe-stepping case	6
c.	Some general points regarding a theory of rational persuasion	8
2.	Reasons	10
a.	Why-questions and their answers	10
b.	Explanatory reasons and normative reasons	11
c.	The explanation condition on normative reasons	13
d.	The ontology of reasons	14
3.	Internalism about Reasons	16
a.	A necessary condition	16
b.	A basic interpretational question	21
4.	The Importance of Internalism for the Theory of Rational Persuasion	24
5.	Two and a half Kinds of Objection to Internalism	25
a.	Kantianism	25
b.	Objectivism	26
c.	Intersubjectivism	27
III.	Objectivism and Epistemic Reason-Giving	29
1.	The Main Argument for Internalism	29
2.	Objections: Rationality and Conversion	34
3.	Objections: Rationality and Responding to Reasons	39
a.	The normativity of internal reasons	39
b.	Williams's challenge and strongly agent-relative aspects of a situation	42

c.	Responding to reasons and the sources of normativity	50
d.	Long-term prudence	53
4.	The Argument as it Stands	54
5.	An Internalist Conception of Epistemic Reason-Giving	55
a.	Epistemic reason-giving.....	56
b.	Advice and teaching.....	57
c.	Freedom and theory	62
d.	Bluffing and blaming	65
e.	Between force and persuasion.....	68
IV.	Externalism and Robust Reason-Giving	70
1.	The Problem with Robust Reason-Giving	70
2.	A Reductive Solution	71
3.	Internalist Answers	75
a.	Straightforward answers	75
b.	Kantian approaches	80
4.	Intersubjectivist Objections to Internalism	84
a.	Desire-independent reasons	85
b.	Empathy and commitment	88
c.	Conditions for a good externalist account of robust reason-giving	91
5.	An Externalist Account of Robust Reason-Giving	92
a.	External reasons	92
b.	Meeting Williams's challenge	93
c.	Meeting the robust reason-giving challenge	99
6.	Robust Reason-Giving and Rational Persuasion	101
V.	Conclusion.....	103
	References	104
	Appendix	110

I. Introduction

This thesis is concerned with reasons for action and the communicative practices in which they play a decisive role. I will try to explain in depth what each is. Right now, I only want to introduce them in showing how pervasive both are in our life. We constantly consider what would be a reason for doing something. We do so both in the public and in the private sphere. In the public sphere, we invoke reasons for or against raising taxes on income, for or against prohibiting smoking in restaurants. In the private sphere, we think about which reasons would speak in favour of or against taking up a certain profession, or which reasons favour going out tonight and which reasons favour staying at home.

Luckily, in all of this thinking we are not alone. Economists try to persuade us that raising income taxes is not a good idea because it will lead to a reduction of investments and economic growth. Other economists tell us that it is a wonderful idea because the state will redistribute the money to people with lower income who will spend more of it and stimulate economic growth through consumption. And, even in our private life, there is a whole cast of others, parents, siblings, friends, teachers, etc., who try to convince us that a certain profession does not fit our personality, or that it will do us good if we go out and meet people instead of staying at home on the couch. It is these practices of thinking about what to do and telling others what to do that are the object of this thesis.

I have two main things that I want to say in this text. The first thing is that a necessary condition on what it is to be a reason is both interesting and holds for a class of reasons which define one type of convincing others to do something. The second thing is that this necessary condition does not hold for another class of reasons which define a different type of convincing others to do something. I will argue for each of these two theses separately and extensively.

The necessary condition to which I referred is a condition proposed by Bernard Williams in a series of papers starting with the seminal piece “Internal and external reasons” (Williams 1979a). Very roughly, it holds that all reasons for action a given agent has derive from her motivational states. Such reasons are internal because they are internal to the agent – they depend on a psychological state of the agent and are thus, in a way, *in* her. In the following, I will call the thesis that this is a necessary condition and that, consequently, all reasons are internal “internalism” or the “internal reason thesis”. The thesis that this is not a necessary condition on what it is to be a reason and that, consequently, there are “external reasons” will be called

“externalism”. What all of this means exactly will be explored in detail later. It should be mentioned right from the start that Williams’s is not the only internalism about reasons that has been proposed.¹ It has been the most influential though, which is why I will concentrate on it.

The theses of this text can therefore be put as follows: (I) internalism is true for one class of reasons which define one way of convincing others to do something. (II) Internalism is false for a different class of reasons which define another way of convincing others to do something.

The method of this thesis consists mostly in constructing arguments building on the existing and extensive literature on internalism regarding reasons and on reasons more generally. For this, it will sometimes be required to interpret important texts in the literature, particularly Williams’s articles on internalism, in order to see both what the theses put forward in these texts mean precisely and how the arguments for these theses work. Nevertheless, this text is not exegetical but systematic. I will look at other texts solely for the purpose of finding out what argumentative routes can and should be taken in order to find out which account of reasons is best.

Some limitations of this project should be mentioned right from the start. Even though I will often speak of a “theory of rational persuasion”, it is not the purpose of this thesis, and it could not be, to supply such a theory. All I want to do here is find some building blocks for such a theory. Also, this thesis will not present a comprehensive doctrine of what it is to be a reason although it may sometimes seem so. I will only deal with one possibly necessary condition on what it is to be a reason and I will ask if and when it is appropriate to demand that it must be met for something to be a reason. Finally, it should be mentioned right from the outset that only the arguments for the first thesis of this work, i.e. that there is one class of reasons for which internalism is true and which define one way of convincing others to do something, may be considered to “show” that the thesis they support is true in a weak sense. By „in a weak sense“, I mean that it claims to show this to people who are already inclined to accept the premises on which the respective arguments rest. The arguments put forward in the last part in support of the thesis that there is another class of reasons which defines a different way of convincing others are only suggestive. By this I mean that they do not claim to consider enough possible

¹ For an overview of different internalisms cf. Darwall 1992; Finlay/M. Schroeder 2017, 1.1. For recent alternatives see M. Schroeder 2007; Markovits 2014, espec. ch. 3.

alternative routes and counter-arguments for counting as a comprehensive treatment of the matter. The second thesis may therefore be regarded as a sustained proposition rather than as a definite conclusion.

The text will be divided in three large parts. The first part is introductory. It has the purpose of setting out the phenomena which are of concern in this thesis in more detail and of introducing the reader to the main positions widely accepted in the literature, which provide the context for the arguments proposed in later parts. The second part will be dedicated to arguing for the first thesis mentioned above. The third part will be dedicated to arguing for the second thesis.

II. Internalism and Rational Persuasion

The main objective of this part is to define what I want to talk about in this thesis. The first two chapters will do so by giving a preliminary description of the phenomena that are of interest here: rational persuasion and reasons. The remaining three chapters will do so by setting out the theoretical debate regarding these phenomena with which we will be engaging in the following parts: the dispute between internalists and externalists about reasons.

1. Giving Reasons

It is a pervasive feature of human existence, in societies at least, that we want to get each other to do certain things. During a pandemic, some people will want to make others get vaccinated; facing the threat of global warming, some people will try to get others to reduce their emission of greenhouse gases; and faced with a man who has threatened to harm his wife, (almost) all people will want to prevent him from doing so. There are various ways of making others do things. For instance, one way of preventing a man from harming his wife is to lock him up in a small room where he is deprived of the possibility to hurt anyone. But, luckily, there are other ways of stopping him from doing this, e.g. taking the keys for the, formerly shared, apartment from him, obligating him to take anti-aggression training and more.²

In this thesis, I want to focus on one specific way of making others do things: what we can approach as giving others a reason to do something. Consider meeting a person who has, ultimately unjustified, doubts about getting vaccinated against a dangerous virus. You may, correctly, tell them that the vaccine is not dangerous and that getting vaccinated will contribute to saving the life of many other people. That the vaccine will save the life of others is a reason for getting vaccinated and in saying that it will, we may say, you “give” that reason to the other person.

For the sake of having a name for what we have approached as giving reasons, and following a remark by Williams (1993, 158), I will call such speech acts “rational persuasion”. I will therefore be concerned with what we may call the “theory of rational persuasion”, i.e. a unified account of such acts that shows us what they consist of and how they are special. This theory might itself be part of a larger theory of how we get each other to do certain things. This larger theory could be called the theory of “power” in something like Foucault’s sense of the term (cf. Foucault 1984, 890). I don’t want to insist on any specific way of drawing connections though.

² See for the legal possibilities in Austria, Deixler-Hübner/Mitgutsch 2007, 93-104.

All I want to say here in relation to this larger theory is that I take threats, or more generally, speech acts which clearly force someone else to do something to belong to the theory of power but not to the theory of rational persuasion even though they are sometimes understood to be reason-giving acts as well (cf. Forst 2015, 64). The reasons for this should become clear shortly.

I think it is obvious why it is important to have a good understanding of what rational persuasion is. Evidently, we give each other reasons to do things all the time. This goes for personal life as much as for politics.³ Moreover, there have been theories that explain and normatively endorse some form of rational discourse as a basic feature of modern societies (cf. fundamentally Habermas 1981a; 1981b; 1992).⁴ It would be too much to ask of this thesis to draw substantive connections here. It is important though that a theory of rational persuasion will have to engage with accounts such as Habermas's and this is a way in which the thoughts presented here would have to be developed further.⁵

In this chapter I want to characterize the phenomenon of rational persuasion. That is, I want to describe some basic features rational persuasion *seems* to have. Those features are, first, that it has some sort of normative standing in that we generally think of it as something good. The second feature is that rational persuasion is not monolithic. Rather, there seem to be two types of it. The *theory* of rational persuasion, with which I will be concerned later, should give an account of these phenomena in that it both explains why they appear to us as they do and in that it shows us whether these appearances are misleading or not.

a. Force and persuasion

There seems to be a deep normative difference between two ways of making people do things: forcing them to do something and convincing them to do something. Consider the example of vaccination again. Whereas it is generally recognised that forcing people to get a vaccine by imposing some sort of penalty is something that is in need of very good justification, and that physically forcing them to do this is hardly defensible at all, attempts at convincing others to get the vaccine by truthfully informing them of its effects are (almost always) not regarded as

³ Perhaps, this way of relating to one another is less relevant in economics, insofar as it is organized by a media like money (cf. Habermas 1981b, 272f.). I don't want to insist on this.

⁴ For an evaluation of one part of Habermas's project see Williams 2002, 225ff.

⁵ Cf. Williams 1996b, 117.

being in need of justification at all.⁶ Contrary to the first, such acts are, in some sense, considered to be a good thing to do. It is these latter acts which I have wanted to single out as acts of “rational persuasion” above.⁷

This is the first element of our phenomenology of rational persuasion. Giving someone reason to do something seems to be a desirable way to get others to do something as opposed to forcing them to do it. This gives us a first desideratum for a theory of rational persuasion. It should tell us why rational persuasion is good, whereas force is, in principle, bad. Or, alternatively, it should tell us that this difference is not as deep or clear as it seems to be.

That the initial appearance may be misleading is suggested by examples of making others do things which cannot be clearly relegated to either the side of force or of rational persuasion. For instance, there is formally truthful but exaggerated marketing for some product or saying things like “men don’t cry”.⁸ This is surely different from physical force, and even threats, in many respects but there must be some kind of difference between this and correctly telling someone that a certain behaviour of theirs will save lives. Regarding these cases, the theory of rational persuasion should either tell us that they really belong to one side rather than to the other, or it should show us how they make the difference between force and rational persuasion collapse altogether.⁹

b. Two types of reason-giving and the toe-stepping case

The second element of our “phenomenology” of rational persuasion is that there seems to be more than one way to rationally persuade someone to do something. Following David Enoch, I will introduce two apparently distinct types of reason-giving: “epistemic reason-giving” and “robust reason-giving” (cf. Enoch 2011, 2).¹⁰ They shall be introduced by giving an example of each.

First, epistemic reason-giving: Suppose I want to tell a colleague what I think of her and that this is not very nice. Informing you about this plan, you might tell me that this is a bad idea

⁶ Cf. Habermas (1981a, 47) on the “forceless force (“zwangloser Zwang”) of the better argument”.

⁷ I will henceforth use the verbs “to convince” and “to persuade rationally” interchangeably. This is not supposed to be a claim about what it is to convince somebody but there seems to be some natural congruity with this idea and what I introduced as rational persuasion.

⁸ See also the description of the injurious power of naming by Judith Butler (1997, 29ff.).

⁹ See III.5.b. and c. for an attempt to account for the apparent goodness of rational persuasion and of different borderline cases.

¹⁰ See also Enoch 2014, 298ff. More on his discussion will be said below (IV.1. and 2.). Enoch also introduces “triggering reason-giving” in his papers. I will say more on this in the next section.

because it will create a bad atmosphere in the department and it will hinder my work and everyone else's. In saying this, you will have given me a reason not to tell that colleague what I think of her. But this reason was there independently of your "giving" it to me. You just indicated it to me or informed me of it. The role of your reason-giving was thus only "epistemic" (cf. Enoch 2011, 3f.).

Second, robust reason-giving: At least some of the people reading (earlier drafts of) this thesis have not done so because they are enthusiastic to learn more about reasons and rational persuasion and they hope to find interesting answers to their questions in this text. They have done so because I asked them to do it. This request on my part has given them a reason to read the thesis in a way that appears to be different from epistemic reason-giving. It does not indicate that there is an independent reason to read my thesis (I asked them to read it precisely because I didn't think it was good enough yet to really be worth reading). Rather, it seems as though the request to read the thesis *was*, or created, the reason to read it for those people (cf. Enoch 2011, 5).¹¹

Stephen Darwall also draws the distinction between epistemic and robust reason-giving.¹² He succeeds to put these two into one example which makes the distinction particularly clear.

Imagine someone steps on your foot (cf. for the following with some modifications Darwall 2013, 136f.).¹³ There are two things you might say to such a person. First, you could say something like "excuse me, you are standing on my foot". Second, you could say "get off my foot please". These two statements do not necessarily have the same meaning. In the first statement you are merely alerting the other person to the fact that she is standing on your foot. Assuming that she recognises that it is bad to stand on someone's foot, she has reason to get off your foot and she will act in order to put an end to this state of affairs, i.e. she will move off your foot. But this reason existed independently of your alerting her to get off your foot. It is only that the world would be a better place if she weren't on your foot and your statement alerts her to the fact that she is standing on your foot.

When you say "please get off my foot", you do more than just alert. You are directly addressing the other person and demanding that she get off your foot. Irrespective of whether she already recognises that the world would be a better place if she weren't standing on your foot, you give

¹¹ See also Herstein (2017, 3110): "directive-reasons" are "reasons to do as one directs because one directed it".

¹² Cf. Darwall 2006, 10. He thereby wants to imply a specific account of robust reason-giving as well – that they rest on "second personal reasons" (for the definition cf. Darwall 2006, 6 and below, IV.3.b.). I separate these two aspects of the example and introduce Darwall's account of robust reason-giving only later.

¹³ See also Darwall 2006, 5-7.

her a further reason to get off: that you demand (or request, depending on your tone) her to do so.

It is easy to see how this corresponds to the concepts of epistemic and robust reason-giving above. Alerting someone to the existence of a reason is a way of informing them about it. If one does this, one only provides help in the epistemic task of finding out what one has reason to do. Demanding someone to do something seems to be one type of creating a reason for them to do something in speaking. The demand itself seems to *be* the reason for doing what is demanded.

It seems, then, that there are two fundamental types of reason-giving speech acts. We have found some ways of intuitively distinguishing them but no principled way of saying what the difference between them is precisely. This is the second desideratum for a theory of rational persuasion. It should tell us precisely what is involved in epistemically and robustly giving reasons and thereby show us what distinguishes them. Or, else, it should show us that the distinction is really not as deep as it seems to be.

c. Some general points regarding a theory of rational persuasion

There are two more things I want to say about rational persuasion right away which may provide a starting point for a theory of rational persuasion. First, it is intimately connected with communicative relations between different persons. Second, it is intimately connected with reasons.

I have introduced rational persuasion as the practice of making others do things by giving them (a) reason(s) for some course of action. This amounts to beginning a theory of rational persuasion by assuming there to be an equivalence between the intuitive idea of rationally persuading, or convincing, others to do something to giving them reasons for action. Indeed, it seems to me that all cases that we intuitively consider to be cases of rational persuasion involve either epistemic or robust reason-giving. This, however, does not imply that all acts of giving someone reason to do something count as rationally persuading the other. This is demonstrated by considering a third type of reason-giving: what David Enoch (2011, 2) calls “triggering reason-giving”.

Consider the following example. I love bananas and, consequently, I eat lots of them. I buy them at my neighbourhood grocer. At one point, she raises the price of bananas significantly. Given that I don’t have an infinite amount of money and that the other grocery stores are inconveniently far away from my apartment, this rise in price gives me a reason to eat less bananas. It doesn’t really create that reason though. Rather, I had a reason all along to adjust my consumption of bananas to their price such that if the price rises I have reason to lower my

consumption. All the grocer does then, is make the conditional in the previous sentence true: she makes the price rise (cf. Enoch 2011, 4f.).¹⁴

The grocer is here, clearly, not convincing me to reduce my consumption of bananas in the sense that we associated with the idea of rationally persuading someone to do something. Rather, it seems, she is just altering the circumstances of my prospective action of buying bananas such that it is less rational. Triggering reasons is hence not even something that must be done by agents. A rainstorm may trigger my reason to stay inside and drink tea if the weather outside is frightful in much the same way in which the grocer triggered my reason to consume less bananas if they are more expensive.

Therefore, we need to devise a necessary condition for reason-giving acts to count as rational persuasion. I will propose one such condition using the theory of speech acts. First, it seems, rational persuasion is something we generally do with words or another system of language. All acts of reason-giving that do not even involve a speech act, such as the rainstorm triggering my reason to stay inside, will be ruled out by this. Second, the speech act must itself be tightly connected to its being an act of giving someone a reason. Using Austin's terminology, we may say that only speech acts¹⁵ which have as part of their conditions of "felicity" (Austin 1962, Lect. II) that they give the addressee a reason for action may count as rationally persuading the addressee. The grocer's raising the price would then not count as rationally persuading me to lower my consumption of bananas. Her raising the price would be just as successful if I had enough money not to care about the price of bananas.

This does not straightforwardly rule out all cases which we may want to exclude from counting as rational persuasion. Threats, for instance, seem to have as part of their felicity conditions that they give the addressee a reason to do something. I propose that they are still ruled out because they may be analysed as altering the practical situation of the addressee by introducing the probability of harmful action into it and simultaneously informing the addressee of this with the intention of thereby giving them a reason for action.¹⁶ Only the latter part would be communicative and would be the concern of the theory of rational persuasion. I don't want to insist

¹⁴ See also Enoch 2014, 299.

¹⁵ Here, I mean what has been called the "illocutionary act" (see Austin 1962, 99f.; see also Searle 1969, 22-25). For a connection of speech acts and giving reasons from a linguistic perspective see Kissine 2013.

¹⁶ Enoch (2011, 5f.) proposes a very similar account of threats.

on this. But, a rough characterization of the reason-giving acts that count as rational persuasion will suffice for now.¹⁷

The second point I want to make here is perhaps rather obvious. An act that is characterized as “giving reasons” must be suitably related to an actual reason. If what is purported to be given in an act of reason-giving (indicated in epistemic reason-giving; created in robust reason-giving) is not, in fact, a reason, the alleged act of reason-giving cannot really count as the giving of a reason and, hence, not as rational persuasion. If it weren’t true that taking a vaccine saves lives, telling someone that it does would certainly not count as (epistemically) giving them a reason to get the vaccine because a fact that does not obtain could hardly constitute a reason for action.¹⁸

In order to specify this condition further, it is necessary to have a good understanding of what reasons for action are. Part of such an understanding would include the question of whether internalism about reasons is true. Answering this question and drawing the conclusions of the answer for the theory of rational persuasion is the goal of this master thesis. However, I will begin with saying a few things about reasons more generally in the next chapter.

2. Reasons

We began by characterizing rational persuasion as giving reasons and by setting out some forms of this giving. I now want to turn to what is given: reasons. I will begin by quickly describing what we talk about when we talk about reasons and then go on to review some of the main points about reasons that are widely undisputed in the literature.

a. Why-questions and their answers

Reasons, or statements involving reasons, provide answers to why-questions. Examples include questions children famously like to ask such as “Why is the sky blue?” and questions adults like to ask such as “Why is the train late (again)?”. These questions ask for an explanation of some fact or phenomenon. As answers to such questions, reasons may provide such an explanation.¹⁹ Not all questions including “why” ask for explanations of this kind though. If you tell your friend that they should watch a certain show, they might ask “Why should I do that?”. The answer to this will also invoke reasons – reasons for watching the show. They are what we may

¹⁷ For more on threats see IV.2.

¹⁸ For more on what constitutes a reason for action see II.2.d.

¹⁹ See on explanatory reasons of this kind e.g. Dancy 2000, 5f.; Alvarez 2010, 27f.

call practical reasons, or reasons for action. These reasons may make your friend see that watching the show would be a worthwhile thing to do. Moreover, your friend may ask “Why should I believe that this is a good show?”. The answer to this will invoke epistemic reasons. If it is a good answer, it might make your friend believe that the show you recommended is a good show.

In this thesis, we will be concerned only with reasons for action. So, the reasons that we will speak about are only those that could come up in a sensible answer to why-questions about actions such as the “Why should I do that?” question. This doesn’t exclude explanations as such from the set of our concerns though. One important thing we can do with reasons (for action) is answer questions of the form “Why did you do that?”. Such questions for explanations seem to be different in some respects from plain questions of the form “Why is this so?”, such as “Why is the sky blue?”. We will see later (II.2.b.) what it is precisely that the “Why did you do that?” kind of questions asks for.

This being said, it is obvious that reasons do not only come up in answers to (actual) questions. Let’s suppose that your answer to your friend’s question about why she should watch the show is: “Because it is so beautifully sad.” This consideration is not only spoken about in the answer. It will also have a role in her thinking about what to do later that evening or in her justifying her watching the show instead of doing something else, like working, to someone criticising her (including herself). Also, it may allow someone wondering about why your friend would spend her evening in front of a screen understand her behaviour. Given her reason to watch the show, her doing so “makes sense”.

b. Explanatory reasons and normative reasons

There is one distinction that pervades the literature on reasons generally: the distinction between normative and motivating reasons (cf. e.g. Parfit 1997, 99; Dancy 2000, 1-4).²⁰ There is still some terminological confusion here as sometimes explanatory reasons are understood to be equivalent to motivating reasons (e.g. Beaulieu 2013, 441f.), and sometimes they are added as a third kind of reasons (e.g. Raz 1975, 15f.; Alvarez 2010, 27). In order to keep the work to do here as minimal as possible, I will only introduce a distinction between normative and explanatory reasons explicitly.²¹

²⁰ See also the overview in Cunningham 2021, 2 and the historical discussion in Dancy 2000, 20-25.

²¹ Dancy 2000, 7 prefers to speak of motivating rather than of explanatory reasons because in a way normative reasons (they explain why an action is right) and reasons that explain why the sky is blue explain something too.

Consider the famous gin-toxic example (Williams 1979a, 102).²² A man is at a bar and orders a gin-tonic. Through some accident, the glass he is served does not contain what he wished to order but gasoline instead. Suppose, he drinks it. If someone asked, baffled about how one could drink gasoline, why our man did this, we may truly answer that he did it because he thought that there was gin-tonic in it. This is what is called an *explanatory reason*. It explains the action by saying what the point of it was to the agent. This is different from explanations of the type of the explanation for why the sky is blue. We say why the man drank what was in the glass by citing something that was appealing to him about drinking it – that he thought the glass contained gin-tonic and that he wanted to have gin-tonic. This notion of explanation was famously named “rationalization” by Donald Davidson (1963, 3).

What we may also say about this situation though is that this man has, or had, a reason not to drink what is, or was, in the glass, namely that what it contains is gasoline. Since he did not know this and as he has already drunk from the glass, this reason has not and will not get to be cited for explaining an action of this man – at least not his not drinking.²³ Rather, it tells us what he should do, or, given that he has already drunk from the glass, what he should have done. This is what is called a *normative reason*. It is what we may have cited if our man had asked us whether he should drink from the glass.

Normative reasons have what has been called “normative force” (Parfit 2011a, 35; Williams 1995a, 36). This has been rendered by metaphors such as that they “count in favour of” (Parfit 2011a, 31) an action, or that they “command, oblige, recommend, or guide” us to do something (Korsgaard 1996a, 8). Other philosophers have tried to give us a better idea of it by saying that the normativity of reasons for action is their property of making some action right or good in some respect (Alvarez 2010, 13ff.).

There is a dispute about whether giving such, largely tautological,²⁴ descriptions is all we can do with the normativity of reasons or whether there is a way to explain it. People who say that we cannot explain it any further argue that the concept of a normative reason is “indefinable”

I nevertheless speak of explanatory reasons here because this is what Williams speaks about (1979a, 102; 2001, 93) and therefore it will minimise terminological confusions in later parts.

²² I slightly modified it to make it more colourful. See Parfit 2011a, 34 for a structurally equivalent example.

²³ It may explain other things such as his regret after having had the gasoline.

²⁴ See Parfit 2011a, 31. John Broome (2004, 41) argues that to say that reasons count in favour of some action is a non-tautological definition of (“pro tanto”) reasons. See Brunero 2013 on this debate.

(Parfit 2011a, 31; Scanlon 1998, 17).²⁵ Others have tried to give a substantive account of the “source” of normativity (see Korsgaard 1996a). We will come back to this discussion later on (III.3.c.).

It should be clear that the reasons that are given in rational persuasion are exclusively normative reasons. When we make someone do something by giving them a reason to do it we tell them what they really should do. We do not invoke explanatory reasons to try to make sense of some behaviour of theirs that “seems strange, alien, outré [or] pointless” (Davidson 1963, 10). We will therefore mainly be concerned with normative reasons here.

c. The explanation condition on normative reasons

Normative and explanatory reasons are not entirely unconnected though. Most importantly, there is a constraint on normative reasons that ties them quite closely to explanatory reasons: the *explanation condition* (cf. Sobel 2001, 145ff.).²⁶ Roughly put, it says that if anything is to be a normative reason to do something for an agent, it should be able to explain the agent’s doing that action. Or, as Williams (1979a, 102) puts it: “[i]f there are reasons for action, it must be that people sometimes act for those reasons, and if they do, their reasons must figure in some correct explanation of their action”.²⁷ Christine Korsgaard (1986, 315f.) has shown that this is something that has been implicitly accepted by almost all moral philosophers of the last two centuries.

Korsgaard has also influentially investigated this condition further. She puts the matter in terms of motivation. According to her, what the “internalism requirement”, i.e., for us, the explanation condition, requires “is that rational considerations succeed in motivating us insofar as we are rational” (Korsgaard 1986, 321). This means that if something is a reason, it will, to some extent, motivate an, at least partly, *rational* person to act accordingly.²⁸ This is a claim both about

²⁵ See also Scanlon 2014, 2: “truths about reasons are not reducible to [...] non-normative truths [...] nor can they be explained in terms of notions of rationality”.

²⁶ Korsgaard (1986, 317) calls it the “internalism requirement”. Since it is different from internalism about reasons, I will rather follow Sobel’s terminology in order to avoid confusion. Dancy proposes a parallel condition on explanatory, or motivating, reasons: the “normative constraint”: “motivating reasons should be the right sort of thing to be normative reasons” (Dancy 2000, 103).

²⁷ For a formalised statement of this condition see Paakunainen 2018, 144.

²⁸ What is at issue here is being partly rational and not being completely rational because not all reasons are decisive. It may be that, although there is a reason for one course of action, there are stronger reasons for another course of action and, all things considered, it is therefore more rational to do that second thing. On different strengths of reasons see Raz 1975, 25-28. In Broome’s (2004) language I am speaking about “pro tanto reasons” here and in most of what follows. Differentiating between this and other kinds of reasons, such as “exclusionary reasons” (Raz 1975, 35-48) is a complication I wish to avoid in this thesis.

reasons and about the psychology of rational agents. If we find out that something is a reason, we will know that rational beings are (psychologically) such that they might act for those reasons. Conversely, though, if there is no way a rational being might act according to a consideration in favour of some action, there is no way it could be a reason for her. Korsgaard and Kant may be thought to press the first aspect (Korsgaard 1986, 330), Williams and Hume may be thought to press the second (cf. III.1.).²⁹

As a necessary condition on what may count as a reason, the explanation condition therefore gets a large part³⁰ of its force from the concept of a rational person. Considerations³¹ which seem like potential candidates for meriting the title of a reason for action will only be excluded from that title by the explanation condition if a *rational person* would not be motivated by them at all³² and they could therefore not be cited as an explanation of this agent's action. Hence, the constraining force of this condition depends on what it is to be a rational person.³³ Much of the discussion regarding internalism will focus on this issue (see below III.1.).

d. The ontology of reasons

The last section has been concerned with a necessary condition on something to count as a reason for action. What has been missing in this is a clarification of the something to which this condition applies. We need to look for the basic ontological substrate of reasons. That is, we should inquire into what larger kind of thing it is that could be a reason if it meets some further conditions, one of which is the explanation condition.

Basically, there have been four contenders for this ontological substrate: facts or states of affairs (Dancy 2000 ch. 3; Alvarez 2010, 40ff.); psychological states (Mitova 2016) and considerations or propositions (Morganti/Tanyi 2017).³⁴ The mainstream view today is that normative reasons

²⁹ I will not discuss Hume's place in all this. Williams does connect his discussion to his position (cf. Williams 1979a, 102) but both Williams himself (1996b, 114) and Korsgaard (1986, 325f.) emphasise the differences. For a broader discussion of Williams's relation to Hume see Sagar 2014.

³⁰ It may get some of the force from the concept of explanation that is at issue as well. This will prove to be more important in the debate on internalism. See below III.2.

³¹ Or whatever reasons most basically are. The next section will be concerned with this question.

³² This addition is, again, due to the fact that not all reasons are decisive. See fn. 28 above.

³³ This is roughly the thesis of Korsgaard's paper: arguments that practical reason alone cannot motivate us to act one way rather than another ("motivational skepticism" about practical reason) have to proceed from a substantive view about what practical rationality is and what it is not ("content skepticism" about practical reason; Korsgaard 1986, 311).

³⁴ Setiya (2014, 221, n. 2) considers true considerations to be equivalent to facts.

are facts. Ulrike Heuer (2012, 223) even says that there is probably no one today who would deny this.³⁵

It is not possible to engage substantially with this debate here. For this reason, I will accept the mainstream position, i.e. that normative reasons are facts, for the purposes of this thesis. Indeed, I am convinced that explanatory reasons, too, are facts, or appearances of facts.³⁶ Accordingly, when I speak about the ontological substrate of reasons in the following, I will speak of facts. Nothing crucial in later arguments should depend on this view.³⁷

Two more things that are important for later parts should be added here though. First, reasons are what comes up as the objects of attitudes which we might describe as reasoning. They are what is considered when we think about what to do and not the act of considering what to do itself (Dancy 2000, 52; 107).³⁸ More precisely, reasons for action may be described as that which may come up in an agent's thinking about how to act as speaking in favour of some course of action (cf. Setiya 2014, 221-223).³⁹ The attitude of practical reasoning, in turn, may just be described as considering reasons for or against some course of action. Sometimes, this is presented as a tentative argument for the view that reasons are facts (cf. Alvarez 2010, 42). For our purposes though, it is not important whether these arguments are sound.

It follows from this that when we are speaking about reasons for action, we are implicitly taking over the perspective of someone considering what to do. It follows that both citing normative reasons for some course of action and explaining action in terms of reasons presupposes putting oneself in the attitude of reasoning of an agent faced with the practical situation in question. We will speak about different ways to imagine oneself to be in the practical situation of someone else later (see below III.3.b.).

Second, the claim that the ontological substrate of reasons are facts implies only that all reasons are facts. It does not imply that all facts are reasons. There may be facts that speak in favour of no action (consider the fact that there are 371 leaves on the tree outside my window) and that

³⁵ On the ontological substrate of motivating/explanatory reasons see, for instance, Cunningham 2021: Here, psychologism, i.e. the view that reasons are psychological states, is much more widespread.

³⁶ For an argument that all reasons (explanatory and normative) are facts, see Dancy 2000, Ch. 5. For the notion of "apparent reasons", see Alvarez 2010, 140 and Parfit 2011a, 35; 111.

³⁷ It may be the case that psychologism about normative reasons implies internalism about reasons. I cannot consider this argument for internalism in this thesis. It seems to have much stronger presuppositions than Williams's argument, which will be considered here, though.

³⁸ Alvarez (2010, 45) calls something similar to this an "act/object" distinction.

³⁹ I am not claiming that this description is not tautological (cf. Parfit 2011a, 31; above II.2.b.).

thus lack “normative force” (II.2.b.).⁴⁰ Moreover, any fact that does speak in favour of an action must satisfy at least the explanation condition if it is to count as a reason, i.e. it must have the ability to motivate a rational agent and, consequently, to come up in a correct explanation of what this agent did (see II.2.c.). I will interpret internalism as proposing such a necessary condition too. Hence, it does not follow from the assumption that reasons are facts that internalism is false. Internalism may just be understood to claim that reasons are facts that speak in favour of an action and entertain a specific relation to the agent’s motivational psychology.

3. Internalism about Reasons

It is now time to set out the thesis on which the debate that is the object of this thesis has centered: internalism about reasons. I will proceed by considering and interpreting a classic statement of the thesis by Bernard Williams. I will then move on to say something about the basic question of how to interpret it.

a. A necessary condition

In a paper that follows up the initial “Internal and external reasons”, Williams puts the internal reason thesis as follows:

A has a reason to X only if he could reach the conclusion to X by a sound deliberative route from the motivations he already has. The externalist view is that this is not a necessary condition, and that it can be true of A that he has a reason to X even though A has no motivation in his motivational set that could either directly, or by some extension through sound deliberation, lead him to X. (Williams 1995a, 35)⁴¹

Let us unpack what is in this dense quote. First, it presents a necessary condition on someone’s having a reason.⁴² If one accepts this condition, one is an internalist, if one does not accept it and holds that there are instances of A having a reason to X that do not satisfy it, one is an

⁴⁰ Alvarez (2010, 42f.) holds the view that all facts are reasons “by virtue of being potential premises in (theoretical or practical) reasoning”. Besides her speaking about both theoretical and practical reasons, this view need not be in substantive disagreement with the view advocated here. From an internalist perspective, for instance, it could be accepted if Alvarez is taken to be speaking about a broader concept of “reasons” than we usually do which includes facts that *could* satisfy the explanation condition and that *could* entertain the connection to an agent’s motivational set that internalism deems necessary. Indeed, there could be an agent with the desire to cut down trees with 371 leaves on them.

⁴¹ I have replaced ϕ by X in order to use the same variables throughout this text.

⁴² Williams (1995a, 35f.) considers it to be a sufficient condition as well. That it is one is not part of the debate on internalism however so I will not discuss it here. On the question to what this necessary condition applies precisely, see below II.3.b.

externalist. Second, the necessary condition contains, in essence, four elements: the *(i) conclusion to X* must be reachable *(ii) for A* by *(iii) sound deliberation* from *(iv) his actual motivations*, i.e. the ones that he currently has. I will say some words about each of them.

The first element will not receive much attention in this thesis. We can simply take the conclusion to X to be the outcome of practical reasoning, whatever that might be exactly – one plausible candidate would be the intention to X (but see Raz 2015 for criticism of this view).

The second element, that it is reachable for A, seems not to say much more than that the agent must also be the subject of deliberation about what to do. This hardly seems controversial or prone to misunderstanding. I will therefore not say much about it either.

The most interesting parts of this condition are certainly (iii) and (iv). Much of the subsequent discussion of how to understand internalism has focused on these two and Williams too has felt the need to specify them further. We should therefore spend some time clarifying what is meant here.

Williams introduces the notion of sound deliberation by providing a list of things that count as such. He wanted to be very liberal about it and included not only simple means-end reasoning but also thinking about how to combine the satisfaction of some motivations or which of one's projects has most weight, "or, again, finding constitutive solutions, such as deciding what would make for an entertaining evening, granted that one wants entertainment." (Williams 1979a, 104). Notably, imagination "about what it would be like if [some development] came about" (Williams 1979a, 105) has a role to play here too. You may thus come to the conclusion that there really is no normative reason for you to go to a party by imagining how exhausting it will be for you.⁴³

There is something more that must be included in the notion of sound deliberation, however. Arguably, there is no amount of imagination, finding constitutive solutions, or other thinking that could have made it clear to the man in the gin-toxic example that, really, he should not drink what is in the glass. The man would only have drawn the conclusion not to drink what is in the glass if he had come to learn the truth about the relevant facts, i.e. what is in the glass. Coming to learn the truth about the relevant facts must therefore be part of what it is to deliberate soundly in the necessary condition imposed by internalism (cf. Williams 1995a, 36).

⁴³ On the vagueness of these notions see Williams 1995a, 38. This is not a problem because it really often is vague what we have reason to do.

In a later paper, Williams argues that we can build this into the notion of sound deliberation because (almost)⁴⁴ every rational agent wants to make rational and well-informed decisions (Williams 1995a, 37). What is meant thereby is not that rational agents want to be well-informed as a matter of fact. Rather, wanting to be well-informed is a transcendental condition of sound deliberation (cf. Williams 1985, 64). Just like you can't sensibly ask "Why should I reflect?" because answering this question would imply reflecting (cf. Williams 1985, 124), you can't sensibly ask "Why should I try to be well-informed?" because answering this question would imply trying to know the facts relevant to its answer. Trying to be well-informed is necessarily part of being a rationally reflective agent and this is why it can be included in the notion of sound deliberation (cf. Williams 1996b, 111).⁴⁵ We may conjecture that anything that a rationally reflective agent necessarily wants would be included as well (this will prove to be more important regarding Kantian theories, see II.5.a.).

It should not come as a surprise then that there are two notions of sound deliberation. One, only slightly idealised, where the rational agent does not get everything she necessarily wants and therefore may be ignorant of some relevant facts. This is what we may call being "rational". The man in the gin-toxic case is probably rational on this standard. The other, more highly idealised, where we include in the notion of sound deliberation that the agent would get everything she necessarily wants. We may call this being "reasonable", just to keep the two apart.⁴⁶ In this sense, the man in the gin-toxic case is not reasonable for he does not know everything that he necessarily wants to know.⁴⁷ This latter notion of sound deliberation is the one referred to in the necessary condition imposed on reasons by internalism.⁴⁸

Let's turn to motivation then. When speaking about motivation, Williams tells us, he has in mind something like desires but also "dispositions of evaluation, patterns of emotional reaction,

⁴⁴ See Williams 2001, 92. That we don't always want to be well-informed can be illustrated by the peculiar situation in which someone does not want to be rational (e.g. when having fallen in love). In a way, then, they might have reason not to be rational. Indeed, it is not crazy to assume that we all have reason not to be too rational (see Schraubroek 2012, 61-63).

⁴⁵ See also Finlay 2009, 15.

⁴⁶ This has nothing to do with the Rawlsian notion of "reasonableness" (Rawls 1993, 48-54).

⁴⁷ On the difference between these two notions see e.g. Williams 1996b, 113; Setiya 2004, 271f.; Parfit 2011a, 34-37.

⁴⁸ From this, it follows that we may call an agent who acts contrary to what she has reason to do unreasonable but not irrational. This is something that Williams admitted as a reaction to the objections by McDowell and Scanlon (cf. Williams 2001, 93; 1995b, 192; 1996b, 109).

personal loyalties, and various projects, as they may be abstractly called, embodying commitments of the agent” (Williams 1979a, 105).⁴⁹ The set of these states is what Williams calls the “subjective motivational set” of the agent, or, for short, her “S” (Williams 1979a, 105).

Motivations can be understood according to a wide concept of what has been called a “pro-attitude” (Davidson 1963, 3f.). Davidson (1963, 4) characterizes pro-attitudes most generally as “attitudes directed toward actions of a certain kind”. As such we can, most widely, understand them as dispositions for action that are actualised if a number of conditions are met. Finlay (2009, 3) calls this the “maximally inclusive set”.⁵⁰ I will refer to this concept as the “dispositional” concept of motivation. Being motivated to help others in need then just means that one will help others if they are in need and there is nothing else holding one back. A number of commentators have objected to this understanding that it renders internalism entirely void (cf. most prominently Schueler 1995, 53-57; cf. also Searle 2001, 215; 217; see below III.1.).

Different interpretations are available. For one, Williams has sometimes been understood to be talking about desires that involve more than simply the disposition to carry out some action in that they are “seeking *satisfaction* of their representational content” (Finlay 2009, 3 with further references).⁵¹ This view has hardly any support in the text however. Williams expressly mentions that he even takes beliefs to be viable candidates for constituting motivations. (cf. Williams 1979a, 107).⁵² Williams thereby rejects the “Humean theory of motivation” according to which motivation must always include at least one desire or a desire-like state (cf. Finlay 2009, 3; 20).⁵³ In what follows, I will take Williams’s position on this for granted.

This need not leave us with the dispositional understanding though. In what follows, I will propose a third understanding of motivation which I take to be Williams’s: what I will call the “phenomenological” concept of motivation. Williams (1979a, 103) interestingly describes elements of S as “deliberative question[s]” to which some course of action provides an answer.

⁴⁹ The term “commitment” is ambiguous. As becomes clear in a later paper where he does not regard someone’s being married as a reason for treating their partner nicely if they have no motivation whatsoever to do so (cf. Williams 1995a, 39), Williams does not mean commitments as they might be understood on the model of promises to someone else (for this understanding see e.g. Gilbert 2018, ch. 8 and 9). Rather, what he seems to mean is a commitment one has as a sort of motivation understood on the model of someone’s being genuinely committed to further some cause. Searle (2001, 169) criticises Williams for not distinguishing commitments from other items in this list but without making that difference.

⁵⁰ On analyses of desires as dispositions see T. Schroeder 2015, 1.1.

⁵¹ Cursive in the original.

⁵² Another locus classicus regarding motivating beliefs in Williams’s work is his notion of „thick ethical concepts“, cf. Williams 1985, 143f.

⁵³ On the Humean theory of motivation see Rosati 2016, 3.1.

This indicates an interpretation of motivation that is oriented by the way an agent having that motivation experiences it. An agent may experience a desire as the question of how to satisfy it and values she endorses as the practical problem of how to promote them. Motivations generally may thus be interpreted as practical problems that call for some course of action to resolve them. I experience the fear of failure in my studies as the problem of how to pass a test or the value of honesty as the question what an honest person would do on some occasion.

The difference between the dispositional and the phenomenological concepts of motivation is clearest in the case of unconscious motivations. An unconscious fear of my mother's harsh judgment may be interpreted as a disposition for action, e.g. subordinate behaviour towards people who are like her in important respects or conformity to what I think she would tell me to do. It does not, however, present itself as a deliberative question for me, nor for a dark part of myself: the unconscious.⁵⁴ On the dispositional concept, this will count as a motivation to act. On the phenomenological concept, it does not.

There is evidence that Williams uses the phenomenological concept. He somewhat obscurely says that in order for rational reflection to derive a course of action from it, it must be conscious. Otherwise the relation between motivation and the conclusion to act would be merely "symbolic" (Williams 1979a, 103).⁵⁵ He thereby excludes those motivations that are only dispositions for action at least from the subset of S that is relevant to sound deliberation and therefore to internalism. I therefore conclude that we should think of motivations in the phenomenological sense as practical problems. This should be expected for the use of the concept of motivation in a condition on what reasons are, given that reasons are the objects of attitudes of reasoning (see above II.2.d.).

We can now summarise what Williams means by the concept "sound deliberation from current motivations" and hence what internalism amounts to. Sound deliberation is a sort of reflection

⁵⁴ Such an interpretation of the unconscious would make it into a full-fledged person which is unacceptable (cf. Lear 2005, ch.1).

⁵⁵ Cf. Finlay 2009, 16f. Elsewhere Williams writes that there may be a sound deliberative route from a motivation to the conclusion to act even though "unconscious obstacles" inhibit the agent from actually perceiving it (cf. Williams 1995b, 188). This is because sound deliberation may be considered to include the removal of such obstacles although there may be some vagueness here – something that Williams says is actually so and should be reflected by theory (cf. Williams 1995b, 188f.). The basic thought at issue here is whether the removal of unconscious obstacles or the discovery of unconscious desires is, like the discovery of the truth about relevant facts, something that any rational agent must want solely in virtue of being someone rationally deliberating. The answer is not important here. But it shows that even if unconscious desires were relevant for sound deliberation, it would not be because they are dispositions for action but because their becoming conscious is included in the problems that necessarily pose themselves to any thinking being.

that is idealised insofar as it includes correctly carrying out some activities that count as rationally deliberating and that the agent achieves everything she must want to achieve (notably, full information about the relevant facts) in virtue of being engaged in reflection. Motivations are the mental states that, first-personally lived through, pose practical questions or problems. Practical deliberation can therefore be understood as the project of finding answers or solutions to them.⁵⁶ Reasons are facts that come up in such practical deliberation as speaking for one course of action (see above II.2.d.). Internalism is the view that a fact is a reason for the agent to do X only if it speaks in favour of action in idealised deliberation starting from the practical problems posed by the current motivations of the agent. Such a fact is an internal reason. A fact that speaks in favour of an action without being connected to the agent's practical problems defined by her current motivations is an external reason. Therefore, there are no external reasons according to internalism.

b. A basic interpretational question

The preceding analysis of the necessary condition imposed by internalism has been in an important way ambiguous. We have not yet specified clearly *on what* it is a necessary condition. In the quote analysed above (II.3.a.), Williams says that “A *has a reason* to X only if ...”, thereby implying that the condition is on the state of having a reason. In the original paper, he says that he wants to talk about internalist and externalist “interpretations” of the “*statement*” “A has a reason to X” (Williams 1979a, 101). This has given rise to two fundamentally different ways of interpreting internalism.

The difference between those two ways of interpreting it can be rendered as a difference in what people have taken the necessary condition evoked above to be about: (a) it is about what a normative reason for action is and therefore an *ontological thesis* (e.g. Milgram 1996; Setiya 2004; Scanlon 1998, 365); (b) it is about the statement “A has a reason to X” and therefore a *conceptual thesis* (Finlay 2009, 13ff.; Parfit 2011b, 269ff.; Skorupski 2007, 79).⁵⁷

Stephen Finlay criticised the ontological interpretation, preferred by most commentators (see Finlay 2009, 2ff.), in his immensely rich paper “The Obscurity of Internal Reasons”. He not only points to the various places where Williams himself has made it clear that he only wants to talk about different interpretations of the statement “A has a reason to X” (cf. Finlay 2009, 13), he also retraces the criticisms that have been brought forth against Williams's argument

⁵⁶ For a similar conception of practical deliberation see Raz 2015, 73.

⁵⁷ On Parfit's conceptual interpretation see below III.3.a.

(Finlay 2009, 2ff.) and seeks to show that they are based on a misunderstanding of it (Finlay 2009, 8ff.; 20). I will come back to Finlay's interpretation of Williams's argument later (see below III.2.). Here, I only want to see whether some sort of ontological interpretation of Williams's thesis may still be defensible.

Surely, Finlay is right to some extent. Williams clearly says that it is not just that, as a matter of fact, there are no external reasons. It's that it is not clear what people are saying when they say that someone "has a reason to X" and take this phrase to have the external meaning, i.e. they take it to imply nothing about the motivations of the agent (see Williams 1979a, 109). Nevertheless, there is a way in which one may say that this is a claim not only about the statement "A has a reason to X" but about what reasons are.

As Julia Markovits (2017, 57) has pointed out, there are two ways one could understand internalism to be a thesis about what reasons are.⁵⁸ One is to think of it as a "first-order normative thesis", i.e. as a thesis which concerns what reasons someone really has. The other would be to consider it as a "meta-normative thesis", i.e. as a thesis that concerns what it is to be a reason. Just as it is part of what a triangle is that it has three edges, it may be part of what a reason is that it bears some relation to the agent's motivation.⁵⁹

Understood as a meta-normative thesis, internalism may be usefully spoken of both as an ontological thesis and as a conceptual thesis. As an ontological meta-normative thesis, it is about what it is to be a reason. As a conceptual thesis, it is about what the concept of a reason picks out. Presumably, the concept of a reason should capture precisely what it is to be a reason. I therefore assume here that the meta-normative and the conceptual interpretation of the thesis are close enough to be treated as equivalents.⁶⁰ Both take internalism to say something about what we could intelligibly hold to be a reason and not only about what reasons for action there are, as a matter of fact (or, rather, moral, or otherwise normative, principle). I will therefore sometimes speak of internalism as a thesis about what it is to be a reason and sometimes I will speak of it as a thesis about the concept "reason". I do not intend to say different things thereby.

⁵⁸ Markovits speaks about subjectivism about reasons. There are various ways of differentiating internalism and subjectivism (cf. Paakkunainen 2018, note 8). They are surely very similar though. For one interpretation of the difference – Derek Parfit's – see below III.3.b.

⁵⁹ But note that this is not "what a reason is" (see e.g. Wiland 2012, 31 on the difference between Davidson's and Williams's treatment of reasons) just like three edges are not what a triangle is. A triangle is a geometrical figure. A reason is a fact (see above II.2.d.).

⁶⁰ Whether they are depends on whether specifying the meaning of a word is equivalent to defining the essence of the object to which it refers (see Fine 1994, 10-14). I cannot go into this here.

There is another reason for thinking that we should understand the necessary condition internalism imposes to be about the statement “A has a reason to X”. As Finlay (2006, 15; 2009, 15; but see also Finlay 2009, 17) also points out, Williams did acknowledge that people might say that “there is a reason for A to X” and that this may mean something entirely different from “A has a reason to X” (Williams 1985, 213).⁶¹ As Williams sees it, this would just mean that we, who say that there is a reason for A to X, think that it would be a good thing if A did X (cf. Williams 1996b, 109).

This is an entirely different concept of reasons than the one we have discussed until now. Consider the following two statements: (1) anyone who is playing chess has a reason not to sacrifice her queen, i.e. that it is the strongest figure; (2) Sarah, who is playing chess with her little nephew, has no reason not to sacrifice her queen. All she wants to do, is make the game fun for her nephew and for this it is not relevant whether she loses her strongest figure. These two statements do not conflict because they use different concepts of a reason. The second statement understands reasons to be facts that count in favour of some action for the agent in the practical situation in question. The first statement understands reasons to be facts that count in favour of some action for someone who endorses what Finlay (2006, 8) has called a “system of ends”. In this case, the system of ends in question is constituted by the end which someone playing chess would normally have, i.e. to win the game.

In this thesis, we will only be concerned with the concept of reason used in statement (2). It is also the one that is more important. Finlay (2006, 5; 7ff.) has described reasons as they are understood in statement (1) as reasons that “don’t matter”. This is because they may neither serve to guide anyone’s practical deliberation nor will they help us in explaining anyone’s action. If you tell Sarah, for instance, that she should not move her queen in this way because it will be captured by her nephew and she will thus lose her strongest figure, Sarah will be right in replying that this is no reason for her not to move her queen in that way.⁶² The concept of reason which is used in statement (1) is therefore unimportant in the sense that it does not have any of the functions we have attributed to reasons above (II.2.a.).⁶³ Not considering it in a discussion on what reasons are will thus not be harmful.

⁶¹ In the original paper, Williams does not make that difference, see Williams 1979a, 101. On the difference between “there is a reason for A to X” and “A has a reason to X”, see also Raz 1975, 20f.

⁶² On differences in the appropriateness of using reasons relating to these systems of ends in conversations, see Finlay 2006, 12f.

⁶³ See Teichmann 2019, 216f. for a similar argument against doubling our way of speaking about reasons.

When we use the phrase “there is reason for A to X” in the sense attributed to Williams above – meaning roughly, “it would be good if A did X” – the system of ends implied by this statement might be constituted by the ends we think a moral agent would endorse. It therefore uses the concept of reason that is implied by statement (1) and is of no interest to us here. It is important to note that this does not show that external reasons just drop out of the set of concern of this master thesis. An externalist may argue that Williams makes a mistake in associating our moral reasons with the “reasons” implicit in the game of chess. Contrary to the latter, moral reasons ought to come up in an agent’s deliberation about what to do in a given practical situation.⁶⁴ Therefore, the externalist may continue, they really do count in favour of an action for an agent. Whether this is correct or not will be discussed in the second part.

4. The Importance of Internalism for the Theory of Rational Persuasion

Before moving on to considering whether internalism is correct, I want to pause to connect internalism to rational persuasion. Above (II.1.), we have proposed analysing rational persuasion as one person giving another reason for doing something. We have inferred from this that it is necessary for a speech act to count as an act of rational persuasion that what it purports to give is actually a reason (II.1.c.). It follows that if a speech act “gives” something that is not a reason, it is not an act of rational persuasion. As internalism imposes a constraint on what a fact must be like in order to be a reason for action, it will rule out some speech acts from counting as rational persuasion.

This may be illustrated in using the example Williams uses in his original paper to argue against externalists: Owen Wingrave’s decision of whether he should join the military or not (cf. Williams 1979a, 106ff.). Owen’s family says to Owen that he should join the army because it would further family honour. It thereby purports to (epistemically) give Owen a reason to do so: the fact that, if he would join the army, family honour would be furthered. Owen, however, does not care for family honour. He has no motivation that is connected with it in such a way that sound deliberation might lead him to join the army. Therefore, this fact does not satisfy the condition internalism imposes for a fact to be a reason.⁶⁵

⁶⁴ Cf. Parfit’s (2011a, 144f.) two concepts of normativity.

⁶⁵ Saying that it is a reason because Owen’s family has a motivation that would be served by Owen’s action (cf. Finlay 2006, 15f.) is misleading. The fact that Owen’s joining the army will further family honour is a reason for them – indeed, it is their reason for telling Owen to join the army. But this is not the reason that they purport to give Owen. They purport to give Owen a reason for Owen’s action of joining the army and not for telling Owen to join the army. For those who are worried, it is no more a problem that one fact may constitute more than one reason than that one hall may constitute the location for more than one event.

Given the above, it follows that, if internalism is true, Owen's family has not given Owen a reason to join the army and their speech act does not constitute an act of rational persuasion. If internalism is false, we may not draw this conclusion. Whether the fact that Owen's family cites as a reason for Owen to join the military really is a reason will be an issue of first-order normative theory. Even though this is doubtful in this case,⁶⁶ we could not decide the issue before getting down to the business of settling what kinds of considerations are important and what kinds of things should be done. Whether internalism is true or not thus has precise and important consequences for a theory of rational persuasion.

5. Two and a half Kinds of Objection to Internalism

In this last chapter of the first part, I want to give an overview of the objections that have been raised against internalism in order to give the reader an idea of what will follow in the next parts. I will begin by considering a Kantian reply to internalism and I will try to show why Kantian theories are incompatible with internalism only at first sight. Then, I will move on to introduce the most prominent objections to internalism defending a more objective notion of reasons ("objectivism"). Finally, I will introduce a different line of objections according to which internalism is not sufficiently attentive to our social nature ("intersubjectivism").⁶⁷ As both deny that reasons are related to the motivations of the agent as suggested by internalism, both intersubjectivism and objectivism are a species of externalism.

a. Kantianism

According to Kantian approaches, what the agent has reason to do does not, at least not solely, depend on what desires she has. Rather, reason, or rationality, alone may determine what we have reason to do. This does not violate the explanation condition because reason may motivate the agent, according to Kant, through a special emotion named "reverence for the law" (Kant 1788, A 130ff.).⁶⁸ The content of this reverence is clearly determined by rationality though. Thus, it may seem as if internalism, in Williams's formulation at least, is false according to

⁶⁶ Cf. Milgram (1996, 205) who points out that Williams's choice of example is therefore not ideal.

⁶⁷ Williams (2001, 94f.) divides the objections to internalism into similar categories.

⁶⁸ "Achtung für das Gesetz" in the original. It should be noted that elsewhere Kant introduces more "subjective conditions for the receptivity for the moral law" (Kant 1797, A 35) including notably "conscience" ("Gewissen") (Kant 1797, A 37-39). There might be interesting connections to draw between this and Williams's discussion of blame (see below III.5.d.). Unfortunately, I cannot discuss this here.

Kantian views. It is rather that rational deliberation determines our motivations than that motivations “control” (Williams 1979a, 104) rational deliberation (cf. Korsgaard 1986, especially 328).

The idea, however, that being motivated to do some things is involved in the very activity of rationally deliberating is not alien to internalism. We have seen above (II.3.a.) that someone rationally thinking about what to do must want to be well-informed. Likewise, Kantian views hold that someone rationally thinking about what to do just could not fail to be motivated to observe the categorical imperative. This is aptly expressed by Kant himself in saying that the test of whether a maxim is permissible is whether it “*could* be willed as a universal law” (Kant 1786, B 52, 77).⁶⁹ According to Kantians, the other option would be self-contradiction which is not supportable for a rationally thinking being. Insofar as it is rational, this being must want not to contradict itself. This is a motivation in the internalist sense. One that any rational being has qua being rational.

Williams (1995a, 44, n. 3; 2001, 93f.) therefore recognises Kantian theories as a “limiting version” of internalism. Of course, they lose much of its characteristic bite. What we have reason to do will no longer depend in such a large part on the contingent and characteristic content of the motivations of the agent because everyone will have the same motivation, implied in their being rational, to comply with the categorical imperative. But this motivation is still a motivation of these agents (cf. Williams 1996b, 111). It still presents a practical problem to them and consequently, if an agent has a reason to do something derived from the categorical imperative, she could arrive at the conclusion to act accordingly by soundly deliberating from her actual motivations. This is all that internalism demands.

b. Objectivism

Objectivist objections hold that reasons do not depend on anyone’s actual motivations. Usually, this comes with an emphasis on the normative nature of reasons which is supposed to be incompatible with their depending on the actual psychological make-up of an agent (cf. Parfit 2011b, 285f.).⁷⁰ Reasons should be able to guide our actions and for this they must be respected

⁶⁹ My emphasis. I altered the translation slightly to fit it better into the text.

⁷⁰ See also McDowell 1995, 77 charging Williams of “psychologism”.

by a rational being.⁷¹ It seems that this could only be the case if they do not depend on what we, more or less contingently, want.

There are two basic accounts of reasons that are available to the objectivist. Either, they do not depend on some thing's being valued by a subject at all. We will pursue this thought in speaking about Derek Parfit's (2011a, 2011b, 2017) latest work on this issue. Or, they depend on some thing's being valued, where the subject of value is no actual person. This thought is expressed by John McDowell (1995) in a by now famous contribution to the debate. Both of these ideas will be examined more closely in the next part.

Regarding the different kinds of reason-giving (see above II.1.), the debate between internalists and objectivists seems to focus largely on a model of epistemic reason-giving. Roughly, objectivist accounts hold that reasons are out there to be responded to by agents. Therefore, it is to be expected that the prime aspect of reason-giving that they will focus on is indicating to someone that there is a reason to which the agent should respond. Consequently, we will also discuss what an internalist account of epistemic reason-giving may look like in the next part.

c. Intersubjectivism

More or less recently, another type of criticism to internalism has been raised: what I will call "intersubjectivist" objections. Contrary to objectivism, they do not hold that reasons are independent of anyone's actual motivations. Rather, they hold that not only facts which stand in a specific relation to the agent are reasons for her to do something. It suffices if the fact is suitably related to the motivations of other agents to whom the agent stands in a specific relation.

Intersubjectivist objections may be divided according to the relation they think must hold between the agent and the other whose motivations may make some fact be a reason for her. Candidates for this relation include "commitment" (Searle 2001, 167ff.); "recognition" (Peterson/Samuel 2021, 83); "empathy and shared intentionality" (Hourdequin 2012); and "facing the other" (Crowell 2020). In the third part of this thesis, I will review Searle's and Hourdequin's models and, subsequently, propose a model of my own for what this relation may look like. I must confess that my treatment of possible positions in this field will be much less comprehensive than my treatment of objectivist objections to internalism. This is due to the fact that the debate has largely focussed on objectivist objections so far.

⁷¹ Cf. the importance of the image of a guiding light in Plato and later stages of Western metaphysics (cf. Taylor 1989, espec. 121-124) and Parfit's (2011a, 46) use of the image of (sun-)light: "If subjectivism is true, we must make our choices in the dark." See also below III.3.c.

Intersubjectivist objections may be regarded as focusing on a model of robust reason-giving. Other agents are not just conceived as people who indicate to the agent that there are reasons to be observed but they, or their attitudes or actions, are what *is* or *creates* (see above II.1.b.) reasons for the agent. Consequently, the third part will connect the debate between internalism and its intersubjectivist critics to the context of robust reason-giving. More precisely, I will suggest that internalism fails to give a satisfactory account of robust reason-giving and that, therefore, some sort of externalism along the lines of intersubjectivist objections to internalism is correct.

III. Objectivism and Epistemic Reason-Giving

This part has two objectives. First, I will try to reconstruct and assess the debate between Williams's internalism and objectivist externalist objections. I will do this by setting out Williams's main argument and follow the discussion concerning it through. At the end, I hope to convince the reader that the objectivist externalist objections fail and internalism holds against them. Second, I will try to draw the consequences of this discussion for a theory of rational persuasion, and, more specifically, epistemic reason-giving. I will, first, argue that we should think of it on the model of advice. Second, I will try to show that we should think of its distinctive mark, which gives it its normative standing, as furthering the practical freedom of the agent to whom reasons are given. I will conclude by raising some doubts about the clear distinction between force and rational persuasion.

1. The Main Argument for Internalism

I will begin by setting out the simplest interpretation Williams's argument for internalism has received (cf. Milgram 1996, 198; Parfit 1997, 112).⁷² According to this interpretation, the argument proceeds on two very basic conceptual premises:

(P1): Any reason must be able to explain the action for which it is a reason.

(P2): All actions must be explained by reference to the agent's actual motivations.

The first premise is a wide version of the explanation condition (see above II.2.c.). It is wide insofar as it does not include a reference to rationality. The second premise is a very plausible assumption regarding motivation. Even on the phenomenological interpretation, motivations seem to be just the things for which agents may act, and which, consequently, must figure in a correct explanation of the action.⁷³

From these two premises, the conclusion that reasons must involve a reference to motivation is drawn:

(M): Any reason must involve a reference to the agent's actual motivations.

⁷² See also M. Schroeder 2007, 7; Parfit 2011a, 108; Markovits 2014, 29f.; Parfit 2017, 297; Paakkunainen 2018, 151f. The reconstructions of this argument differ in details.

⁷³ On the dispositional interpretations, it is a tautology. (P2) may be a bit too strong in that it neglects other ways of explaining human actions, e.g. by reference to unconscious motives. I use explanation in the narrow sense here in that it makes us see what the agent saw in the action, i.e. it gives explanatory reasons (see above II.2.b. and below III.2. for clarification). In this sense, (P2) seems to be correct as it stands. I am grateful to Martin Niederl for prompting me to clarify this.

The conclusion that internalism is true is obtained by constructing this reference with the help of sound deliberation:

(I): A has a reason to X only if she could arrive at the conclusion to X by soundly deliberating from her actual motivations.

This is justified because it is hardly disputable that there must be a reference to rational thinking in the explication of what it is to be reason.

One objection clearly seems open to the argument from (P1) and (P2) to (M) (cf. Parfit 2011a, 108).⁷⁴ What if a reason brought about a motivation of the agent? It would be conceivable that there is a reason for me to do something and that this reason explains that I believe that there is a reason for me to act. This belief, or a desire brought about simultaneously, could then motivate me to act (cf. McDowell 1995, 70f.).⁷⁵ It is thus not the motivation that partly explains why I have reason to act. Rather, it is the other way around. That I have reason so to act explains why I am motivated so to act.⁷⁶

Williams anticipated this objection. He correctly imagines that the external reason theorist will claim that an external reason could explain A's doing X through A's coming to believe that she has this reason to do X and her coming to be motivated to do X by this belief (cf. Williams 1979a, 107). Consequently, he further elaborates the argument in order to give a reply based on the hardly disputable fact that a reason's ability to explain the behaviour of an agent must have something to do with the agent's being rational.

In order to appreciate the force of this reply, we must have one more look at the explanation condition. In any formulation, it includes a reference to an *ability*. What an ability is precisely is notoriously hard to define and seems to depend in a peculiar way on the context of the utterance (cf. Schwarz 2020 for a recent contribution to this debate). In what follows, I want to put matters as simply as possible and only say one thing about abilities generally.

For an entity E to have the ability to do something X, it is not sufficient that it is merely *possible* that E does X (cf. Paakkunainen 2018, 153).⁷⁷ An ability includes the further reference to some

⁷⁴ See also Parfit 1997, 113; Finlay/M. Schroeder 2017, 1.3.

⁷⁵ See also Dancy's (2000, 101) discussion of the "three-part story".

⁷⁶ See also Parfit 1997, 102; Searle 2001, 191.

⁷⁷ Paakkunainen speaks of "capacities". I don't think that she means something different thereby.

changes in E's properties or circumstances that count as the *actualisation* of this ability implying that other changes which also result in E doing X do not count as such. A full account of an ability will include a full account of these actualising changes.⁷⁸

To make this clearer consider the following example. Suppose that I say that I can drive a car. Relative to the context, I might truly say this even though I don't have a car I could drive right now so that the actualisation of my ability to drive would require that I get a car. I could not truly say this, though, if I had not yet acquired the relevant motor skills, knowledge of traffic regulations etc. although I could acquire them. The possibility of these changes to make it the case that I am driving a car does not support my saying that I have the ability to drive a car. Hence, they are not included in the changes that count as actualising my ability to drive a car (cf. Paakkunainen 2018, 153f.; Schwarz 2020, 8-10).

Therefore, in order to understand the explanation condition correctly, we need an account of what changes leading to the reason's explaining the agent's action count as actualising the reason's *ability* to explain the agent's action. Clearly, not all possible changes count. The fact that they sell pistachio ice cream at a specific corner of the city is not able to explain my going there if I hate ice cream. It is not able to explain this despite the fact that it is possible that a confused scientist could alter my brain structure so that I do develop a strong desire for pistachio ice cream (cf. Sobel 2001, 148 for examples of this kind). If this actually happens, of course, the fact that they sell pistachio ice cream somewhere will be able to explain my going there. But this does not mean that it has had that ability before already.

When introducing the explanation condition above, we have, following Korsgaard, already implicitly proposed an account for which changes in the reason's circumstances or properties may count as actualising its ability to explain an action: a normative reason to X must be able to explain the agent's doing X if she is rational, in some respect. The changes that count as actualising the normative reason's ability to explain an action are therefore changes that imply the agent coming to be rational in some respect. Williams can be regarded as embracing this account in saying that the external reason theorist must think of the agent's coming to believe that she has reason to X as being the result of "considering the matter aright" or "deliberat[ing] correctly" (Williams 1979a, 109).⁷⁹

⁷⁸ This is supposed to be a rough and ready account sympathetic to a context-sensitive "quantificational analysis" of abilities according to which "A has the ability to X" should be understood as "A Xs in an accessible possible world" (for more on this and for a defense of this approach see Schwarz 2020).

⁷⁹ "Becoming rational" may sound as if it implied that rationality is not primarily to be understood as a process (a position often attributed to Williams, cf. Finlay 2009, 3 with further references; see also below III.2. regarding

Also, as the gin-toxic example illustrates, what is at issue here is, strictly speaking, not rationality but being reasonable (see above II.3.a.). The change to the agent's circumstances that we imagine in the gin-toxic case is not one of the agent coming to reason better on the grounds of the evidence available but one of being better informed and, more generally, in a better position to make rational decisions.

Consequently, we should alter (P1) such that it reads:

(P1'): Any reason must be able to explain the action for which it is a reason if the agent is reasonable.

The if-clause captures the changes that count as actualising the reason's ability to explain the action in question.

Stated like this, this could hardly be a controversial premise.⁸⁰ But, by itself, it could also not provide the resources for a reply to the externalist objection. This is because such a reply would presuppose that being "reasonable" must be understood as soundly deliberating on the basis of the agent's actual motivations. If this is substituted for "reasonable" in (P1'), external reasons would indeed lack the ability to explain reasons if the agent is reasonable. They are defined as those "reasons" that may not explain the agent's acting according to them given that she deliberates soundly from her actual motivations. The premise (P1') would then contain the part of the conclusion that divides internalists and externalists: the connection of reasons and reasonableness, and the agent's sound deliberation from her actual motivations. This way of arguing would thus clearly be circular (cf. Paakkunainen 2018, 153).

In fact, the concept of being "reasonable" allows for two totally different interpretations. Whereas Williams considers someone to be reasonable, roughly, if she deliberates about the matter in question soundly, externalists may understand "being reasonable" to mean nothing other than recognising the relevant normative reasons and being motivated by them (cf. Parfit 2011a, 62f.).⁸¹ Parfit calls this latter view of rationality "substantive" as opposed to "procedural" (Parfit 2011a, 62; 78).⁸² If this notion of being reasonable is adopted, nothing much

McDowell) but as a state. I do not mean to imply that here. For, people who understand rationality to be primarily the process of thinking rationally may consider starting the activity of soundly deliberating to be the change which is picked out by the word "becoming rational" and which we thus count as a change that actualises a reason's ability to explain action. Thanks to Martin Niederl for prompting this comment.

⁸⁰ Finlay 2009, 3 calls it the „platitude“ of a connection between rationality and reasons.

⁸¹ See Finlay/Schroeder 2017, 1.3; Searle 2001, 170.

⁸² By "procedural" rationality, he means the kind of rationality that we have opposed to being reasonable above though.

changes if (P1) is exchanged for (P1') in the argument. If a reason is part of the explanation of A's doing X, this will just mean, certain other conditions being met, that she is being, perhaps even fully, reasonable. For the externalist, the if-clause in (P1') is basically a tautology.

Some people have considered the possibility that Williams himself might actually be operating with such a substantive notion of rationality. This point has been noted regarding the wide concept of motivation. If reasons should play any role in our lives, humans should be such that they may respond to reasons.⁸³ Therefore, we certainly have some sort of disposition to respond to reasons. If even this disposition is included in the subjective motivational set, Williams may be formally right – all reason statements will include reference at least to this disposition – but the thesis will not say anything distinctive (cf. Setiya 2004, 269).⁸⁴ Certainly, we should avoid attributing this understanding to Williams.⁸⁵

Now it seems to be clear in which direction Williams's reply to the externalist objection would need to go. We may expect Williams to claim that "reasonable" in (P1') must be interpreted in a procedural rather than in a substantive sense. It is here that commentators have been baffled that Williams does not seem to present an argument at all. He just seems to take such a concept of rationality for granted (cf. Finlay 2009, 4; 6 with further references). Rather, his reply to the externalist objection seems to consist in an unrelated conceptual argument (cf. Finlay 2009, 11). Recognising that his opponents will say that an external reason will motivate the agent in virtue of the agent coming to believe that she has that external reason to act, Williams asks "what it is that [the agent] comes to believe" then (Williams 1979a, 109)? But, this point seems to be even weaker (cf. Finlay 2009, 12). The external reason theorist will just say whatever she thinks external reasons are (as McDowell 1995 could do), or that it makes no sense to look for a further analysis of the concept of a reason (cf. Parfit 1997, 121).

Stephen Finlay has convincingly argued that understanding the conceptual argument to be unrelated to the argument with which we started out is a misinterpretation of Williams's argument.

⁸³ Cf. the first sentence in Parfit (2011a, 31).

⁸⁴ See also Schueler 1995, 53-57; Parfit 1997, 118f. Searle's (2001, 217) statement of this point is harder to interpret. Even though he does accept the explanation condition (cf. Searle 2001, 171, point 5), in this statement he seems to deny that an agent would need to have any disposition to be motivated by (desire-independent) reasons. Surely, he would have to make an exception for the disposition to recognise reasons, i.e. what he calls "recognitional rationality" (Searle 2001, 181). Then his tautological version of internalism is equivalent to the one set out here.

⁸⁵ Williams may of course still assume that we have a desire to be reasonable. But, given his notion of "being reasonable", the content of this desire will itself be bound to the motivations of the agent because the reasons the agent has are bound to those motivations and "being reasonable" is defined as responding to the reasons the agent has. The problem therefore goes back to what it is for the agent to "have a reason".

It is the final step of the argument that begins with the explanation condition (cf. Finlay 2009, 19).⁸⁶ Whereas Williams's internalism provides an entirely clear notion of being reasonable that does not involve the concept of "reason" in it, i.e. sound deliberation from existing motivations, the externalist's notion of being reasonable is itself defined in reference to the concept of the agent's reasons. It is the quality of responding to reasons. But this means that it cannot serve to analyse the notion of "reason" on pain of circularity. We are still in need of an account of what it is that the agent responds to when she responds to a reason. If we do not have such an account, we will not know what a reason is and, consequently, we will not know what responding to a reason, i.e. being reasonable, is either.

Williams's argument thus finishes by issuing a challenge to the externalist (what I will call "Williams's challenge" in the following): give a non-circular account both of what reasons are and of what it means to be reasonable such that (P1') and (P2) are true.⁸⁷ Internalism's advantage is that it has no problem at all with giving us such an account. This is why we should accept it.

In what follows, I will look at attempts to respond to this challenge. I will concentrate on the accounts proposed by John McDowell and Derek Parfit. McDowell keeps a central place for deliberation from the motivations of an agent but instead of the actual agent he takes an ideal counterpart of hers as the point of departure. Parfit stays with the actual person's character but he understands deliberation as coming to recognise reasons that she has independently of anyone's character. Let's take each in turn.⁸⁸

2. Objections: Rationality and Conversion

McDowell's (1995) objection is quite easy to trace because he accepts one of the main tenets of Williams's position: that reasons depend on motivation. He proceeds by doubting only Williams's restriction on who the subject of these motivations must be – the actual agent that is

⁸⁶ Finlay's reconstruction of the argument is more complicated than just that. I will come back to parts of it below (III.2.).

⁸⁷ This is point 5 of Searle's (2001, 171) list of things he needs to do with regards to desire-independent reasons.

⁸⁸ There have been many other propositions on what external reasons could be, e.g.: Milgram (1996, 206f.) thinks they might be converting experiences (this might be regarded as somewhat of an intermediary position between McDowell and Parfit); Lubin (2009, 284ff.) thinks they are transcendental conditions of reasoning (a Kantian outlook, Williams would probably think is really internal; cf. Teichmann 2019, 217); Teichmann 2019, 216 thinks that one's motivations themselves need a reason (this too is not something Williams would deny but that seems to be manageable in an internalist framework). I don't have space to discuss all suggestions here. I chose the two accounts that have probably been the most influential and, more importantly, have interesting flaws that help to spell out the idea of Williams's internalism in more detail.

said to have the reason. Instead, he proposes to understand reasons as being relative to the motivations of an ideal agent modelled on the Aristotelian conception of a “properly brought up person” (cf. McDowell 1995, 73).

McDowell’s attack can be framed as a proposal to understand “reasonable” in (P1’) differently from Williams without using the notion of a reason to define it. As we have seen (above, II.3.a.), for Williams, being reasonable means deliberating soundly which involves a set of deliberative procedures carried out in ideal deliberative circumstances, i.e. those circumstances that the agent, qua rational being, must want to obtain. This understanding, however, cannot be presupposed without begging the question against the externalist (see above III.1.). Consequently, Williams uses the much less determinate notion of “considering the matter aright” for what we have called “being reasonable” in his argument. What would therefore be required for the externalist to answer Williams’s challenge is “that the agent should acquire the motivation *because* he comes to believe the reason statement, and that he should do the latter, moreover, because, in some way, he is considering the matter aright” (Williams 1979a, 108f.).⁸⁹

It is here that McDowell steps in. Williams seems to suppose that considering the matter aright is a kind of *activity* that proceeds from the agent’s actual motivations. That is, the agent should come to the conclusion to do what she is supposed to have reason to do *by* some sort of rational deliberation. McDowell accepts that the *activity* of practical deliberation must proceed from the motivations of the agent (cf. McDowell 1995, 74) but he doubts that being rational *only* means to carry out that activity. Rather, we can also understand rationality as a *state* in which a person can be. McDowell defines this state using the idea of “ethical upbringing” understood in a “roughly Aristotelian way” (McDowell 1995, 73).⁹⁰ Being rational, or, rather, reasonable, thus means not only to rationally think about what to do but to do this on the motivational background of a person who has been brought up well.

This concept of rationality as a state can be put to use to define “reasonable” in (P1’), and, with it, the kind of change that may count as the actualisation of the reason’s ability to explain the agent’s action. Williams seems to understand “coming to consider the matter aright” as something that must be attained “*by* deliberating correctly”. Instead, McDowell proposes that we may think of this transition in the agent as a “transition *to*” considering the matter aright

⁸⁹ Emphasis in the original.

⁹⁰ For McDowell, it would be a difficult question what such a “properly brought up person” would be like. Whether there is a consistent answer to this question has no direct bearing on internalism though, as McDowell’s proposal can be rejected for other reasons as well. I will therefore not engage with it any further.

(McDowell 1995, 78).⁹¹ The change that is picked out in (P1') by "if she is reasonable" and which restricts what may count as an actualisation of a reason's ability to explain an action may thus be defined by its *outcome*: the agent being like the properly brought up person, in at least some respect. "Being reasonable" in (P1') therefore means not only carrying out the activity of sound deliberation in ideal circumstances but the *state* of being someone whom we may call reasonable because she is like someone who has been properly brought up. This provides us with an intelligible version of (P1') which does not refer to the concept of a reason and may therefore be helpful in elucidating it.

In this, it should be noted, no reference to intrinsically normative facts "out there" is made (McDowell 1995, 73; 80). It is only a matter of considering an idealised version of the agent. And, as McDowell (1995, 76f.) reminds us, the agent in Williams's notion of sound deliberation is also idealised to a certain degree. At the very least, she knows the truth about all the facts relevant to what she should do.

In his reply to this objection, Williams (1995b, 187) claims that a misunderstanding has arisen. Williams's argument does not presuppose that what defines the condition of actualisation in the explanation condition designated by "reasonable" must be defined solely in terms of the nature of the process, i.e. certain activities that count as sound deliberation, as opposed to the outcome, i.e., according to McDowell, coming to be a properly brought up person. Rather, Williams's notion of sound deliberation itself includes changes defined by their outcome, e.g. that the agent comes to learn the truth about relevant facts or removes "unconscious obstacles" to sound deliberation (cf. Williams 1995b, 188) – these are the changes that the agent must want to happen qua rationally thinking about what to do (see above II.3.a.). For Williams too, part of what it is to be reasonable is to be in a certain state, that of being in ideal deliberative circumstances, and not only to carry out some activity. At first sight, this may seem like a point in favour of McDowell's argument. But, Williams goes on to say how this argument fails to provide a satisfactory analysis of reasons and (P1').

The problem is that on McDowell's account of reasons, the statement "A has a reason to X" doesn't say anything distinctively about A, i.e. the person who has the reason (cf. Williams 1995a, 189). Rather, it is a statement about what a properly brought up person would do if she were in A's situation.⁹² Reason statements would then not "relate actions to persons, but types

⁹¹ My emphasis.

⁹² Admittedly, McDowell does not talk about what a properly brought up person would do but about what A would do if she were properly brought up (cf. the difference between (C) and (G) in Williams's reply (1995b, 189)). It is hard to see, however, how being a sound deliberator could be something different for A than for anyone else

of action to types of circumstances, and they are most revealingly expressed in the form 'in circumstances C, there is reason to X' (Williams 1995a, 190).⁹³ And, this is unacceptable. Statements about the reasons for action someone has must be in a substantive way about that agent.⁹⁴

One may continue to ask why this should be so, though. Indeed, there are two possible answers to this question and both are revealing about the deeper motivations of internalism.

The first answer is given by Williams in his reply to McDowell (cf. Williams 1995b, 190-192). It proceeds on a distinction between the good and the right. The good is what should be done, or, differently put, what a, in all respects, perfect agent would do.⁹⁵ If the right, or what someone has reason to do, is to have any significant content as an idea, it must be different from the good (cf. Williams 1995b, 191f.). It is if it is distinctively about what a specific agent should do and not just about what a perfect agent would do in her situation. Williams's point here is that if we extend the concept of normative reasons and no longer take it to be concerned with what is right for a particular actual agent but with what is right for an ideal agent, it will just collapse into the notion of the good and we will have lost an idea that is important for understanding both ourselves and others.

The second answer is never given directly by Williams. Rather, it has been set out somewhat systematically by Stephen Finlay (2009, 13ff.) in giving a novel interpretation to Williams's argument. This route has the advantage of not looking like a new argument for internalism but incorporating the idea of reason statements being distinctively about the agent into the original argument from the explanation condition. Indeed, it focusses on a word in (P1') that has received much less attention than the idea of "being reasonable": "explain".

In distinguishing explanatory from normative reasons, we have already discerned two concepts of explanation (see above II.2.b.). There are explanations of the type of saying why elephants

without assuming that internalism is true (cf. Williams 1995b, 190f.) and this would be necessary for reason statements in McDowell's sense to be distinctively about A.

⁹³ I changed the variables to fit the use in this text.

⁹⁴ Cf. Williams's insistence that practical reasoning is "radically first-personal" (cf. Williams 1985, 23; 74). When I think about what to do, I think about what would make sense that I do and not, at least not necessarily, about what should be done. See also Taylor 1989, 76 on the connection between one's reasons for action and one's identity.

⁹⁵ These definitions are probably circular but it is not clear that the good can be defined in a non-circular way. On Williams's concept of the good, see Williams 1972, 38ff.; 1985, 65. According to those remarks, the good seems to be the accordance to normative requirements implicit in the concept of a thing. The paradigm case is something like a good house, i.e. one that keeps you safe, warm, etc. (see also Korsgaard 2009, 27-34). The good that I have set out here is what it is to be a good agent or a good human. It should be noted that Williams is generally skeptical of this latter idea (Williams 1972, 55-62).

are so big. And then, there are explanations that tell us why someone did something. The first will probably just invoke causal connections, perhaps along with some structural features of the world. The second invokes an idea of subjective justification. It will allow us to see what the point of the action was to the agent, i.e. how it has seemed to be a good idea to her.⁹⁶ Explanatory reasons, as we have introduced them above, are what we invoke to give this latter kind of explanation.

It should be clear that “explain” in the explanation condition should be taken in the second sense. It is that someone who has a reason must be able to *act for that reason* and therefore to embrace, or follow, the normative force of that reason (cf. Korsgaard 2008, 214).⁹⁷ At least in some possible circumstances, A’s reason to X must give us the subjective justification of her doing X, i.e. it must show us what the point of doing X was for A.

If “explain” is understood in this sense, a normative reason for A to do X must be such that it must show us how doing X could seem to be a good idea to A – if she is reasonable. Of course, given (P1’), the normative reason only needs to have the *ability* to explain actions in this sense. But, this means that the account of the actualisation of its ability (see above III.1.) needs to include some elements about how the normative reason in question could come to show us what the point of the action is for the agent. Otherwise, all the facts that are relevant to assessing whether a fact could provide subjective justification of an action for an agent could change in the actualisation of this fact’s ability to do so. This “ability” would then mean no more than that this fact could explain what the agent saw in the action if the agent were (completely) different. This is surely too substantive a change to count as actualising an ability. If the reason’s ability to explain the agent’s action is to have any significance beyond being the mere possibility of explaining the agent’s action, this actualising change must proceed from some fact about the agent that is true before the change already. At least in this regard, then, a statement about A’s reasons must be distinctively about her.

Any account of what normative reasons are must be able to account for this substantive relativity to an agent.⁹⁸ McDowell’s account cannot do this because, according to him, the actualising

⁹⁶ For expressions of this kind see, again, Davidson 1963, 3.

⁹⁷ For a detailed discussion of “acting for a reason” see Mantel 2018. Unfortunately, I could not extend this research further into this field.

⁹⁸ Cf. for similar arguments Brunero 2007, 34 and Finlay 2009, 14ff. Finlay’s way of putting this point is different at least in that he speaks of “evidence” for an action being a good thing to do in a given situation. I don’t think this is particularly helpful. There are still big differences between practical reasons and evidence – at the very least if something seems to be a good idea to the agent, there is still the moment of intention which mediates between this reasoning and action. Evidence will, most often at least, just make us believe something if we consider it right (cf.

changes of the ability of normative reasons to explain A's actions include such things as "conversion" (McDowell 1995, 74), i.e. a total change of the relevant facts about the agent. Put in the terms Williams uses: on his account, reason statements are not distinctively about the agent to whom the reason is attributed. It therefore fails to give a satisfactory account of what a reason is and of how (P1') should be understood.

I take this latter argument to be crucial against McDowell's externalism. It fails to capture something we just do expect normative reasons to do for us. They are neither just an ideal standard by which we lead our lives, nor are they simply tools for predicting the behaviour of human beings. They are a way of understanding the actions of ourselves and one another in a deeper sense – making sense of actions by regarding them as an expression of what the agent considers to be good, important, desirable, fun, or necessary.

3. Objections: Rationality and Responding to Reasons

The most recent contributions to the debate on the problems connected with internalism about reasons have focussed mostly on Derek Parfit's forceful critique of what he calls subjectivism about reasons.⁹⁹ Parfit advances numerous arguments against subjectivism which are directed against internalism too (I will explain the difference between internalism and subjectivism shortly). I will discuss three points here that represent the basic ideas behind Parfit's attack and are likely to reveal major issues in the debate between internalism and its objectivist critics. I will begin with an argument that targets whether Williams is even talking about the right concept, then I will consider whether Parfit's account can meet Williams's challenge and lastly I will say something about one of Parfit's main arguments that subjectivism, and hence internalism, cannot be true.

a. The normativity of internal reasons

The first argument against internalism is that it is not what it purports to be. Whereas Williams himself¹⁰⁰ and most of the literature¹⁰¹ understand his thesis to concern *normative* reasons, Parfit (2011b, 279) thinks that it does not really do so. His argument for this proceeds in two steps.

Parfit 2011a, 47-50). On Parfit's notion of the difference between epistemic and practical rationality see below, III.3.c.

⁹⁹ Cf. Lin 2020; Shemmer 2019; Darwall 2017; Temkin 2017; Markovits 2017.

¹⁰⁰ Williams 1995a, 36; Williams 1996b, 110f.

¹⁰¹ Cf. e.g. Darwall 2017, 261 with whom Parfit engages in the passages just quoted and Scanlon 1998, 365.

First, he holds that Williams's internalism is not about what reasons there are but that it only claims that the phrase "A has a reason to X" means that A would be motivated to X if he deliberated correctly from his already existing motivations (cf. Parfit 2011b, 270). He calls this "analytical internalism" (Parfit 2017, 293; Parfit 2011a, 72) as opposed to "subjectivism" (see Parfit 2011a, 45f. and below III.3.b.) which he thinks is a substantive thesis. A thesis is a "substantive" thesis about reasons, according to Parfit (cf. 2011b, 275), only if it tells us whether some consideration is a reason for doing something and if it is informative. This is an answer to what we have considered to be the question of what reasons there are. Contrary to this, internalism really is concerned with what it is to be a reason (see above II.3.b.). If Parfit wishes to treat this as a conceptual question, his first point will be correct.

Second, he holds that Williams's concept of a reason is not a normative concept at all and that, therefore, there must be a different concept of a reason which we use to say things about what people really should do. This latter concept is the interesting concept of a reason, i.e. the one we use in thinking about what we ought to do on particular occasions (cf. Parfit 2011b, 287f.). Consequently, Williams would fail to speak about the interesting concept of reasons and we can disregard his challenge.

Evidently, the crux of this argument lies in the assertion that the concept of an internal reason is not really normative. Parfit provides an argument for this by splitting up the parts of an internal reason judgment. What we say when we say that someone has internal reason to do something is that she would be motivated to do it upon ideal practical deliberation given her current motivations. Parfit agrees that this is a normative judgment in that there is a normative element in it – ideal deliberation (Parfit 2011b, 285). But this alone, he argues, cannot make the whole judgment normative. Because it is a judgment about what would *actually* follow from the agent's motivations and since the agent's motivations are in themselves not normative – they are part of what the agent is actually like, not what she should be like – the statement about internal reasons is just as little normative as the statement about what one would actually do in an unjust world. The normative element concerns only the circumstances of a non-normative event. What really would be normative is what the agent morally, or perhaps prudentially, or otherwise ought to do in these circumstances (cf. Parfit 2011b, 286).

The analogy Parfit draws, between what someone would do in an unjust world and in circumstances of ideal deliberation, is not sound. It is based on a distinction between the circumstances of an event and the event itself. The action in an *otherwise* unjust world is thus separated from the unjust world in which it happens and which counts as the situation in which it takes place.

But, in this sense, deliberation is surely not part of the circumstances of an action. The last part of any full piece of practical deliberation is the conclusion to act and this conclusion, even if it does not lead to carrying out the action with necessity,¹⁰² is very closely connected to the action itself. It gives it its point, i.e. what distinguishes it from mere movements.¹⁰³ Surely, the point of someone's doing something isn't just a circumstance of that action in the way that the injustice of the world around one is a circumstance of it.

The normativity of internal reasons can be located at the level of this conclusion to act which supports the action. Insofar as this conclusion is congruent with the conclusion at which the agent would have arrived in ideal deliberation, it is the right conclusion, i.e. what the agent should do. And if the agent carries out that conclusion, she will have done the right thing. In ideal deliberation, where that conclusion is right per definitionem, what one should do is then identical with what one would do.¹⁰⁴ Therefore, what one would do in such a situation is also the right thing to do. And surely, the "right thing to do" is a normative notion. Williams's concept of "reasons" is therefore a normative concept.¹⁰⁵

There is another argument for the normativity of internal reasons that may be attractive here: there are practices in which normative reasons play a large role and which seem to rely on an internal understanding of reasons. One of these practices is advising others. Of course, this argument is only suggestive. There are accounts of advising which understand it as the giving of external reasons¹⁰⁶ and perhaps one might deny that what is at issue in advice are really normative reasons. Indeed, I will go the other way around and I will use the arguments for internalism developed here in proposing a specific account of advice. Nevertheless, a more sophisticated version of this argument may be illuminating.¹⁰⁷

¹⁰² For a locus classicus on weakness of will see Davidson 1969.

¹⁰³ Korsgaard (2008, 227-229) makes this connection between an action and the reason for it very strong but something like this will even hold for views like Raz's (2015) who denies that the conclusion of practical reasoning is an action or an intention – if an intention is then formed on the basis of practical reasoning, it must be closely connected to it even if that connection is not itself one of reasoning.

¹⁰⁴ Temkin 2017, 21 holds that Parfit should accept this but say that this is so because being ideally rational just means responding correctly to reasons. I think that he is right in that Parfit should go this way. The issue then comes down to whether this is a defensible account of ideal rationality. I will turn to this shortly.

¹⁰⁵ Parfit may respond to this that it begs the question against his position because it locates normativity at the level of our judgments of the situation and not at the level of the situation. I believe that this would be a misunderstanding of the argument. I only argue here that a concept of normative reasons which locates the normativity of reasons at the level of the judgment of a situation is still a normative concept in that it captures something which may guide our reflection on what to do. Whether this is the correct way of viewing normativity will be discussed only later. I am grateful to Martin Niederl for prompting this comment.

¹⁰⁶ See Thomson 2001, 48-50 and below III.5.b.

¹⁰⁷ Darwall 2017, 274ff. could be understood to present such an argument.

The concept of an internal reason is therefore suitable for saying what people should do. It is thus a competitor in illuminating our use of the concept of normative reasons. Therefore, it should be taken seriously as should the challenge issued in arguing for it. Parfit then has to say how “reason” and (P1’) are to be understood if he wishes to propose a different account.

b. Williams’s challenge and strongly agent-relative aspects of a situation

Let us now turn to the core of the discussion here. The argument proposed by Williams demands that the externalist give us an account of reasons that both makes us understand what reasons are and how they may explain our actions without referring to sound deliberation from the agent’s actual motivation. In dealing with McDowell’s objection, we have added that this account could only do so if it showed how reasons are specifically related to the person having these reasons such that statements about the agent’s reasons are distinctively about her. More precisely, it needs to show how the reason could explain to us what the (actual) agent might see in the action if she does it.¹⁰⁸

Parfit’s account for this basically consists of three parts. One is that the concept of reasons is indefinable. The second is that all reasons are object-given. The third is that they may explain our actions through our recognition of these reasons. Let’s see what this means exactly.

First, we should look at the concept of a reason. According to Parfit, there is no non-tautological way of saying what a reason is (Parfit 2011a, 31). All we may do is illustrate it in saying that they “speak in favour” of some course of action or give clear examples of what reasons are (Parfit 2011a, 31). We may also say that reasons have “normative force” (Parfit 2011a, 35) but the concept of normativity is also indefinable for Parfit, except, perhaps, in terms of reasons (cf. Parfit 2011a, 109; 144).

The second point concerns the question of what reasons there are, i.e. it concerns to which things the concept “reason” should be applied (cf. Parfit 2011a, 72).¹⁰⁹ According to Parfit, there are two rather general theoretical conceptions of this that precede different substantive conceptions of morality. According to objectivism, all reasons are “object-given”. According to subjectivism, all reasons are “subject-given” (Parfit 2011a, 45). Object-given reasons are provided “by facts about the objects of [...] desires or aims” (Parfit 2011a, 45), i.e. “events” in a large sense (cf. Parfit 2011a, 43f.). Subject-given reasons are “provided by [...] certain facts about what would fulfill [...] our present desires or aims” (Parfit 2011a, 45). In the first case,

¹⁰⁸ Parfit 2017, 294ff. explicitly accepts at least some version of (P1’).

¹⁰⁹ See above II.3.b.

reasons are derived from facts about the things we may want or intend to do. In the second case they are derived from facts about our wanting to do these things.¹¹⁰ Parfit (2011a, 47) holds that objectivism is true and subjectivism is false.¹¹¹

The third point is that the way reasons influence our behaviour is that we “respond” to them. Responding to reasons means to become aware of the facts that give one reason to do something and to become motivated to do this thing through becoming aware of these facts (cf. Parfit 2011a, 32).¹¹² Responding correctly to these reasons is what it is to be reasonable (see above II.3.a.).¹¹³ The main difference between responding to reasons for action and to reasons for desire and belief is that the former includes a moment of voluntary decision and the latter does not (cf. Parfit 2011a, 47-50).¹¹⁴

It should be noted that it is the first point that operates most clearly on the level at which we have been operating here. The points that Parfit calls conceptual determine what it is to be a reason on his account, i.e. an indefinable and intrinsically normative entity (cf. Parfit 2011b, 266f.).¹¹⁵ On an internalist view, on the other hand, there is more to be said about what it is for something to be a reason, i.e. that it is necessary that it could make the agent do that for which it is a reason through sound deliberation from the agent’s existing motivations. The question with which we have to deal now is whether Parfit’s minimal notion of reasons can do all that a notion of reasons needs to do.

At first sight, this account of what a reason is, seems, at least, defensible. After all, we do invoke reasons all the time and hardly anyone, except philosophers, it seems, ever has trouble understanding what is meant when we do. If we want to refute this understanding, we will thus need some sort of argument.

¹¹⁰ Parfit speaks of reasons as deriving or being grounded by facts. In this, he does not intend to oppose people who say that reasons are facts (cf. Parfit 2011b, 280). I do not see how anything depends on this either. In discussing Parfit I will sometimes take over his way of speaking about reasons as being grounded by facts. I do not intend this to make a difference to the view that reasons are facts.

¹¹¹ Indeed, Parfit (2011a, 107) claims that, if subjectivism were true, nothing would matter. See Markovits (2017, 71-74) for a response to this.

¹¹² On questions concerning this responding relation see Temkin 2017, 3-12.

¹¹³ Theories that hold this have been called “recognitionist” theories of practical reason as opposed to “constructivist” theories (cf. Darwall 2006, 291-293). Cf. Korsgaard’s (1996a, 36f.) distinction between “substantive” and “procedural” realism (cited by Darwall 2006, 292).

¹¹⁴ For more on this, see below III.3.c.

¹¹⁵ It is also “non-natural” according to Parfit (2011b, 266f.). I will not pursue questions related to the relation of naturalism and normativity any further here.

In principle, the same will go for the worry that Parfit's story cannot give a satisfying account of (P1'). Williams too acknowledges that beliefs about reasons may be motivating. If some facts are reasons for doing something, we may rightly come to believe that they are reasons for acting and this may happen through considering them right. Considering them right, on this view, is nothing but responding to these reasons and it can give us an account of what "being reasonable" in (P1') is. This is circular – being reasonable has reasons in its definition and comes up in a condition on what reasons must be like – but if we understand what a reason is without further ado, this will not be a problem. "Being reasonable" and "reasons" may just belong to a structure of inter-defined concepts that are irreducible to anything else. Here too, we will need an argument. Indeed, since being reasonable and reasons are inter-defined on this account, it might be the same argument as the one mentioned in the preceding paragraph.

Things are no longer so simple if we consider the requirement, discussed in considering McDowell's objection (above III.2.), that reasons must be related to the agent to whom they are ascribed in a distinctive way. If it is indefinable, it seems, the concept of a reason will not give us any resources for establishing this. We may see whether this impression is correct in considering the arguments brought forth for the necessity of a substantive relation of the agent's reason to the agent in discussing McDowell's view. If they have no force against Parfit, this impression will not be confirmed and we must admit that Parfit can establish such a connection sufficiently well.

First, I will consider the need for a distinction between the right and the good. Parfit himself says that a thing's goodness can be that it implies there to be reasons for action for someone (Parfit 2011a, 38). And a thing's being good, as Williams (1985, 65) also remarks, is a property of this thing that is, to some extent, independent of the agent who judges it to be good. A good hotel, for instance, just is one with comfortable beds, clean rooms and some other stuff. But, as Williams urged in reply to McDowell, goodness and reasons are not related in any straightforward way (see above III.2.). There is a sense in which one can have no reason to stay at a good hotel (Williams 1985, 139). For Parfit, the hotel's goodness would imply there to be reasons for (all) agents to do something, e.g. to stay there. It thus seems as though Parfit cannot allow for an agent's reasons and the good to come apart in the way Williams suggests they do.

The strength of this argument should not be exaggerated. It is not universally held that the difference between the good and the right is a difference between normative judgments that are in a distinctive way relative to an agent and normative judgments that are not. Alvarez (2010, 13), for instance, proposes saying that "good" refers to the rightness of actions whereas "right"

also concerns beliefs and emotions. Against the example of the hotel, one may object that what we think of as a good hotel is just a hotel that is good for most people which may differ from a hotel that is good for a particular person. It is the latter sense of “good hotel” that implies that a person has a reason to stay at the hotel.¹¹⁶

We shall therefore turn to the second argument for the special connection of a reason for action and the agent for whom it is a reason that we have considered in discussing McDowell’s objection. We held that “explain” in (P1’) takes the meaning of explaining in showing what the point of the action is to the agent, i.e. in subjectively justifying the action. As (P1’) speaks of the normative reason having the *ability* to explain an action in that way, we need an account of which changes count as actualising that ability. This account consists partly in coming to be reasonable. If coming to be reasonable would imply that all the facts relevant to judging whether some course of action has a point for the agent change, this ability of a reason hardly seems to have any content besides that it would be *possible* that this reason explain what the agent saw in the action if she were different. This surely is too weak for us to be speaking of an ability (see above III.2.).

In a first attempt, we may propose that what needs to be true of the agent already, in order for a reason to have the ability of explaining an action of hers in showing what she sees in it, is that she has some relevant pro-attitudes towards this action which will be cited in the explanation of the action itself (see Davidson’s analysis, above II.2.b.). Parfit’s notion of normative reasons has no place for such pro-attitudes, or motivations. Such a line of reasoning might have had some force against McDowell, as he agrees that practical reasoning must start from motivations. It will not convince Parfit though. He will just say that this justificatory element of the explanation in terms of reasons derives not from the motivations of the agent but from the reasons she could have for these motivations. What shows us what someone sees in defending the rights of refugees is not her caring for other human beings but the fact that refugees are human beings who suffer, which is the reason for caring for them in the first place (cf. Parfit 2011a, 45).

This response gives us an idea of what Parfit’s account for a normative reason’s ability to explain through subjective justification might be. The change that would count as actualising that ability will be the agent coming to recognise that she has a given reason and act on it. The reason will be the justification, its recognition by the agent makes this justification subjective.

¹¹⁶ For the notion of something being good *for* someone as opposed to something being good in an impartial sense, see Parfit 2011a, 38–42.

Understanding “subjective” any other way from the outset would just be to beg the question against the externalist. From Parfit’s perspective, this may be all there is to say about it.

We may doubt whether this is not just the same thing as McDowell’s saying that the reason will show us what the agent may see in the action if she were different. But, in Parfit’s account, there is a clear reference to what the agent is actually like now. We may hold that she must, in principle, be able to recognise that she has that reason. And, this reason will explain her action if this ability is actualised and if she acts on the reason she has thus recognised. The only change that she needs to undergo is recognising that she has that reason and be motivated by that recognition. She need not otherwise become like an ideal agent type. This is a possible account of the reason’s ability to explain in justifying. Without further argument, we cannot ask for more.

I will now present one argument which would imply that this is not enough and that something stronger must be presupposed about the agent’s relation to the reason if the reason should count as having the ability of explaining an action of the agent.¹¹⁷ The argument has three premises. First, in the situation in which an action takes place, we can distinguish between the aspects that are agent-relative in a strong sense and aspects that are not. Second, reasons are agent-relative in a strong sense. Third, there is no principled way of making this distinction except by referring to the motivations of the agent. It follows that the agent’s reasons must be determined by her motivations.

The idea of a distinction between strongly agent-relative features of a practical situation and not so strongly agent-relative features is present in many things we say. Consider for example: “I would love to be in your place.” Or: “What would you do in my place?”¹¹⁸ These phrases refer to the speaker, or the addressee, taking the other’s place. They do, or are asked to, consider the situation of the other as it would present itself to them rather than to the agent who is actually in that place. It is understood that some features of the situation change with the change of the agent who considers this situation. Or, more to the point, the practical significance of some aspects of the situation changes. Someone, let’s call her Gaby, who loves physical exercise may say to her friend, Fiona, who despises it but who has been signed up with her partner for a cross-

¹¹⁷ Williams might suggest something like this argument when he says that McDowell’s account rather relates reasons to circumstances than to agents (Williams 1995b, 190). I don’t want to insist on interpreting Williams this way which is why I will present the argument as my own. I do think that it is roughly what Williams had in mind though.

¹¹⁸ Though this latter locution is ambiguous in that it can also be used in asking for advice. Then, the situation is considered as it presents itself to the one asking this question and the respondent is only asked to make use of her reasoning capacities on these grounds. Sometimes, this is also expressed in the paradoxical: “What would you do if you were me?”

fit class: “I would love to be in your place.” The significance of having agreed to physical exercise is different for the two. Roughly, for Gaby it is desirable and for Fiona it is not. This is an aspect of the situation that is strongly agent-relative. The fact of the agreement to physical exercise itself is not strongly agent-relative in this sense although it is agent-relative in the sense that Gaby is at least originally included in the description of the fact. It is not strongly agent-relative because it is included in what Fiona imagines to constitute her place when she imagines being in Gaby’s place.

Reasons seem to belong to the strongly agent-relative side. When Gaby pictures herself to be in Fiona’s situation, she considers the fact that she will have the opportunity to do physical exercise soon to be a reason for joy. Fiona does not. Rather, she may regard this fact as a reason for trying to get out of it. This seems to be so not only for situations where what is at issue is a (dis)like of the persons involved. Suppose that Gaby is outraged by activists who bothered her to sign a petition for the well-being of animals. Most likely, she did not regard the fact that some animals are not well as a reason for her to step in and sign the petition.¹¹⁹ She may ask Fiona who seems skeptical regarding her outrage “What would you have done?” Fiona may reply: “I would have signed it.” Clearly, this implies that she does regard the fact that animals are not well as a reason for signing the petition. It is the reason-giving significance of this fact that changes relative to whether the situation is considered from Gaby’s perspective or Fiona’s. Reasons therefore seem to be agent-relative in the strong sense.

The congruence of this idea with the idea of reasons as explanations including subjective justifications is quite clear. The fact that animals are suffering may explain Fiona’s action in situations such as being asked to sign a petition but it could not explain Gaby’s action. Given that Gaby does not care about the well-being of animals, and she does not think that caring is morally or otherwise required of her, that fact could not explain her action in this way. If she did sign the petition, we would be quite baffled and search for further explanations. If Fiona signed the petition, that fact alone will be sufficient for explaining the action.

So, now the question is whether there is a specific way to make the difference between what, regarding a situation, is agent-relative in the strong sense and what is not. We will need something that Fiona “takes with her” in imagining to be in Gaby’s place which allows us to determine that some facts have a different significance for her than for Gaby. Williams can be understood to propose one understanding of this. The strongly agent-relative aspects of a situation

¹¹⁹ Another possibility would be that Gaby does not believe the factual statements of the activists. However, it seems to me that most often this is not the problem people have with being asked to sign a petition.

are determined by the motivations of the agent. When Fiona imagines to take over Gaby's place, she assesses the situation on the background of her desires, her projects, her commitments and what she deems valuable.¹²⁰

It follows that what reasons an agent has is determined by her motivations: Reasons belong to the strongly agent-relative aspects of the situation. Which aspects are strongly agent-relative is determined by the agent's motivations. Hence, what reasons someone has is determined by her motivations.

So far so good. Two questions remain. First, do we even need an account of strongly agent-relative aspects of a situation? Second, can Parfit provide such an account without recourse to the motivations of the agent? If the answer to the first question is yes and the answer to the second question is no, Parfit's account of how reasons are tied to one agent specifically will be inadequate.

The answer to the first question seems to me to be pretty obviously yes. The distinction between strongly agent-relative features of a situation is one that, though implicit in many things we say, is not a notion that we really do use in everyday life. Indeed, even my stylised introduction of it has been rather obscure. If philosophy should do anything, it should give us an account of such phenomena. If we say that we don't need one, this just seems like giving up on doing philosophy in this matter.

Regarding the second question, I think, Parfit has a meaningful option to object to the argument. He may hold that Williams's argument rests on a confusion of two concepts of the distinction between strongly relative aspects of a situation and non-strongly agent-relative aspects.

Accordingly, the first sense of that distinction is given by the distinction between agent-relative and agent-neutral reasons. Agent-relative reasons are such that a full statement of them includes a reference to the agent. The full statement of an agent-neutral reason does not (Löschke 2021, 360; Nagel 1986, 152f.).¹²¹ According to this understanding, the significance of the fact that cross-fit is hard physical exercise is (strongly) agent-relative. The full statement of this fact, if it is to have the normative force of a reason, includes a reference to an agent's preference for

¹²⁰ This does not mean that the conversational context cannot bid to regard some motivations as part of the non-strongly agent-relative parts of the situation. But then, they will count as circumstances of the action rather than as what may explain action in the sense of showing us what the point of the action could be for a specific agent. Such a case may be someone being addicted to smoking and asking someone else what they would do in their situation. Obviously, they are not looking for an answer of the type: "Well, I just wouldn't smoke any more, as I have no desire to do that anyway."

¹²¹ This is not all there is to say about this distinction. See, for instance, the rest of Löschke 2021. But it is all that needs to be said here.

hard physical exercise, one might argue. This is why Fiona will say that if she were in Gaby's position she would happily go and do cross-fit with her partner. The fact that animals are suffering, however, is not a reason that includes reference to any agent. Its normative significance does not depend on the fact that it gives the agent pleasure or that it is in some other way related to the agent. It is, therefore, according to this proposal, not strongly agent-relative.

When Fiona says that, contrary to Gaby, she would sign the petition, she implies a different sense of the distinction between strongly agent-relative and agent-neutral aspects of the situation. She does not want to say that what is special about the situation as it presents itself to her is that, for her, the fact that animals are suffering has the significance of being a reason for signing the petition. She wants to say that, contrary to Gaby, she recognises that this is a reason for signing the petition and that she would have acted accordingly (this might be signalled by an accusing tone implying that the situation would have presented a reason to sign the petition for Gaby as well). In this sense, then, the difference between strongly agent-relative and non-strongly agent-relative aspects of a situation is the difference between the things experienced and the experience of things. A tree may look like a bush to someone with bad eyesight and no glasses. A reason for action may not look like a reason to someone who does not consider the matter correctly – for whatever reasons.

That all reasons are strongly agent-relative, Parfit may proceed, is therefore not true in the same sense in which it supports the argument. It is true in the sense that the experience of reasons is always agent-relative. It is not true in the sense that normative reasons are always agent-relative. Still, a conclusion about normative reasons is drawn from this premise, i.e. that we need an account of normative reasons that explains to us how it could be true that all normative reasons are strongly agent-relative. The externalist therefore does not need to accept that conclusion and the challenge it issues and the argument must fail.

I must confess that I don't see how the internalist could respond to that reply. It is true that an internalist view could accommodate these two senses of aspects of a situation being strongly agent-relative. Agent-relative and agent-neutral reasons might be distinguished according to whether sound deliberation from their underlying motivations includes thinking that it would be good if everybody considered this to be a reason for action. Fiona's tone in saying that she would have signed the petition may signal that she thinks that, if Gaby had deliberated soundly from her actual motivations, she would have seen that she also had a reason to sign the petition. The distinction between the experience of a reason and the reason that is experienced is clearly

not a problem for internalism either.¹²² But, all this does not show that externalism is incapable of handling the phenomena set out here. It only shows that internalism is capable of doing so. This is no more than a stand-off then.

c. Responding to reasons and the sources of normativity

It seems, now, that Parfit's externalism is able to satisfy all the points we demanded of an analysis of reasons. Williams's challenge seems to be met and the argument for internalism seems to be, finally, unsuccessful. This, I think, is still merely an appearance. I want to present three closely related arguments for this claim. The first will proceed on a newly introduced premise. The second will try to show that if Parfit's account is accepted, an intuitively significant difference cannot be made. The third attempts to show that on an externalist account we would need to accept an unacceptable description of our practical situation.¹²³

The premise that needs to be introduced for the first argument is one about normativity. We already said above that normativity, or, more to the point, the experience of it, can be described as some consideration commanding, obliging, recommending, or guiding us (see above II.2.b.; Korsgaard 1996a, 8). We are "pushed" to do something, so to say. For instance, some people feel that they should go to church on Sunday. In such cases, though, it is always possible to step back and ask why. Why should I do that? This is what Christine Korsgaard (1996a, 10) famously calls the "normative question".¹²⁴

The answer to this question calls for normative reasons. The person who feels that she should go to church might think that it lies in the fact that this is tradition, that it is required by her religion, or that God will be pleased if she goes. But, here the normative question can be reiterated: for any of these reasons for going to church, we may ask why this should be a reason for going to church. We may wonder why one should follow the demanding force of tradition of religion or why God's desires should guide our actions.¹²⁵

¹²² Cf. Williams 1979a, 103 on the ways in which one may falsely believe an internal reason statement about herself.

¹²³ None of these arguments are new, I might add, although I might present them differently than they are presented elsewhere (as will become clear, I draw a lot on Korsgaard 1996a for this). They are usually not brought forth in discussions of (Williams's) internalism though. Finlay 2006, 4 is an exception.

¹²⁴ In fact, it is what may be called the genus of what Korsgaard calls the normative question. Korsgaard asks more precisely "what justifies the claims morality makes on us" (1996, 9f.; emphasis omitted). I introduce the larger question here to make the steps of the argument more explicit.

¹²⁵ See, from an objectivist perspective, Nagel 1986, 149.

At this level, there are two possibilities for giving an answer.¹²⁶ The first is to give normative reasons for taking the primary reasons seriously. Our person may think, for instance, that she should follow tradition because it provides her life with stability or because she owes it to her ancestors. Any of these answers may be right but it is easy to see that they raise the normative question once again. At some point then, there can be no such answer to it.

The second possibility is to give reasons which explain why some fact has normative force for us.¹²⁷ Here, the answer will have to involve a reference to what it is to be a normative reason. The answer Korsgaard (1996a, 49f.), and Williams's internalism, gives is that some fact is indeed a reason for someone to do something if she will be motivated to do it after ideal reflection, i.e. what we have analysed as "sound deliberation" (see above III.3.a.). The answer to the question why the fact that tradition requires going to church on Sundays is a reason to do that will then be that the agent still feels obliged to do so after ideally reflecting on whether she should follow tradition or not.

The argument against Parfit's conception of reasons is that, if the concept of a reason is considered to be indefinable and if reasons are considered to be intrinsically normative, there will be no answer to some normative questions. Of course, there may be normative reasons for some facts to be reasons for action. The first kind of answer is therefore also available to Parfit. But at some point such an answer is no longer available. For a utilitarian, for instance, there is no bigger normative principle that is the reason why she should try to increase pleasure. At this point, an internalist could (attempt to) say why the fact that some action will increase pleasure is a reason for doing it. Parfit, however, could not even try to give an answer to this question because such an answer would need to tell us, at least implicitly, what normative reasons are. Because, according to him, such an answer is in principle impossible, Parfit would need to hold that raising the normative question at this point does not even make any sense. This is contrary to the premise we introduced. We said that, regarding normative requirements, it is always possible to step back and ask why we should do something. Hence, Parfit's concept of what it is to be a reason is not satisfactory (cf. for this argument Korsgaard 1996a, 38-40).

The core of the second argument against Parfit's understanding of reasons is that, on his account, there will be no significant difference between practical and theoretical rationality. Thinking about what to do would be, just like thinking about what is the case, trying to find out

¹²⁶ For the following, see also Taylor 1989, 76f. on the distinction between offering "basic reasons" and "articulating the moral point of our actions".

¹²⁷ On the power of reflection in this regard see Williams 1996a, 211f. and Korsgaard 1996b, 251-254.

what the world is like in some regard. It will be trying to find out what facts are also reasons for action. The only connection practical reasoning would have with actions would be that the knowledge of what reasons there are, and of what should be done according to them, could be applied in practice (cf. Korsgaard 1996a, 44). It follows that “human life and action consist in the application of theories, theories about what is right or good” (Korsgaard 1996a, 44). But, surely, it is more than that. For instance, it may be the attempt to work out what we should consider to be valuable by thinking and acting on what we do consider to be valuable.¹²⁸

The third point is that the answer to the normative question must be in a distinctive way about the agent. We have seen above (III.3.b.) that Parfit can avoid arguments for the assertion that reasons must be in a distinctive way about the agent and that this can only be captured by relating them to her motivations. Here, I want to propose another argument for the claim that normative reasons must be distinctively related to the agent in the way internalism requires. Roughly, I want to suggest that normative reasons should not only make it clear why an action would be a good thing to do in my situation, i.e. *for* me. It must also show why it would be a good thing to do from my perspective, i.e. *to* me.¹²⁹

If this were not the case, there is no way we could regard ourselves as the authors of our actions. If my reasons would have nothing to do with what seems good to me, the vantage point from which it should be assessed what I should do would lie outside me. If this were so, I could not consider myself a “source of normativity”¹³⁰ but I would have to consider myself being rather like a machine: the question is not what I value or what would be something that I would do but whether I am functioning properly where what determines what counts as functioning properly is a normative principle outside of myself. We would not have a character but we would either be defective or non-defective. Peterson/Samuel (2021, 84) call this idea of normativity as something outside of human beings “normative alienation”.¹³¹ This cannot be right. Therefore, we need something more than Parfit gives us to account for normativity. Something like reflection.

Drawing these three arguments together, we can conclude: first, we need an explanation of normativity; not out of naturalist concerns but in order to answer a question that comes up in

¹²⁸ I think that this captures roughly the idea of practical reflection that is implicit in Williams’s internalism. Cf. Taylor’s (1989, 72) concept of practical reasoning as advancing in comparative stages.

¹²⁹ See Markovits 2017, 67 for a similar argument.

¹³⁰ See Korsgaard 1996a, Lect. 3 for the Kantian claim that our humanity is the source of normativity.

¹³¹ See also Markovits 2014, 37; 54.

living our lives. Second, on Parfit's concept of normativity, we cannot make a significant distinction between practical rationality and theoretical rationality. Third, if the explanation for normativity did not locate the source of normative reasons somewhere inside us, we would be alienated from the significance facts have for us. Life would be like fulfilling a number of chores rather than building a meaningful whole.

All three of the arguments given proceed on assumptions that an implication of Parfit's concept of what it is to be a reason is problematic. I suppose that an externalist could deny all of these assumptions. Parfit does seem to deny them. He clearly considers practical rationality and theoretical rationality not to be distinguished in the way that is supposed by the second argument (cf. Parfit 2011a, 118; and 47ff.; 420ff.). To the third argument, he might be considered to give a metaphorical response in responding to the charge that for the objectivist reasons are "out there" like the sun: "If subjectivism is true, we must make our choices in the dark" (Parfit 2011a, 46). I know of no explicit counterargument to the additional premise of the first argument. This is why I hold it to be the strongest.

Denying these assumptions would come at a cost, however. What inspires the arguments brought forth here, is the thought that the normative significance of the world and ourselves is the product of permanent questioning, reflection, and reassessment on the basis of new experience. The problem is that Parfit cannot show how permanent questioning of this kind is even an option for us. This is why, I think, we must reject his objection to internalism.

d. Long-term prudence

Before concluding this chapter, I will consider one argument to the effect that internalism can't be true because it has untenable consequences. Parfit advances a number of arguments (cf. Parfit 2011a, 73-107) to the effect that subjectivism can't be true. They concern internalism at least insofar as they draw consequences from subjectivism that could also be drawn from internalism.¹³² I will only consider one such argument Parfit proposes here: the agony argument (cf. Parfit 2011a, 73-82). It seems to me to be the strongest of these arguments. But, as it has no significant structural similarity to the other arguments, I cannot pretend to be dealing with the others as well.

Suppose that someone knows that he will suffer from extreme pain tomorrow if he does not take a pill today, but also that he has no desire whatsoever to avoid future extreme suffering.

¹³² For the relation of subjectivism and internalism according to Parfit see above III.3.b.

On subjectivist (or internalist) views, this person will have no reason to take the pill: there is no sound deliberative route from the existing motivations of the agent to taking the pill. But this seems untenable. Given the horrible pain that awaits him, the agent must certainly have a reason to take the pill. Therefore, at least this reason must be objective and subjectivism cannot be true.¹³³

It seems to be really hard to bite the bullet here and to say that this person does not have a reason to take the pill. But, if we look closer, the argument rests on a mistakenly narrow view of practical reasoning. Parfit himself suggests that what we may be envisaging when we say that this person has a reason is a situation where he remembers what severe pain is like and to be so horrified by this that he will see that he has reason to take the pill (cf. Parfit 2011a, 82). But clearly, remembering sensations, just like imagining ourselves in certain situations (see above II.3.a.), is a perfectly sound piece of practical reasoning. What we really want to say is that such a person, i.e. an agent who does not yet want to avoid this pain but will want to do so when he has imagined that pain, has a reason to take the pill. And this is perfectly compatible with internalism. If this person imagined the pain that she would go through (and imagined it vividly enough, we might vaguely add) and still does not want to take the pill, we might actually say that she really does not have reason to avoid it by taking the pill.¹³⁴ The unease that comes with this may just be due to the fact that it is really hard to imagine such a person.¹³⁵

4. The Argument as it Stands

We can now state Williams's original argument for internalism as it stands and as it has survived the most prominent objections that have been put forth against it.

(P1'): Any reason must be able to explain the action for which it is a reason if the agent is reasonable.

¹³³ Note here that suffering and pain are not necessarily (but see Arpaly/Schroeder 2014, 116ff. for a theory that connects pleasure and desires) identical to aversions (i.e. a desire that something not be the case). It seems that I can want to suffer for various reasons just as much as I can want not to have some pleasure (cf. Parfit 2011a, 52-56).

¹³⁴ This is not the only way of dealing with this sort of argument. Another possibility would be to allow deliberation from future desires to ground reasons. See Shemmer 2019 and Lin 2020 for an overview of recent contributions to this debate. These discussions are significantly more complex than what I have put forward here. I think that some reply of the sort given here is able to deal with Parfit's account although some adjustments may be needed.

¹³⁵ See Williams (1979a, 106) regarding a person who allegedly does not care about his health: really we think that "at some level he *must* want to be well" (emphasis in the original). Perhaps, we will also suspect that something might be in other ways "wrong with her" – something that rather calls for other interventions than rational argument (see Manne 2014, 110f.).

(P2): All actions must be able to be explained by reference to the agent's actual motivations.

Any account of reasons must show how these two things come out as true under it. We have seen that an account that regards reasons as the practical conclusions an ideal agent type would reach through sound deliberation is inadequate (above III.2.) because it cannot show us how the reason is able to explain an action by the agent for whom it is a reason and not by someone else. An account that holds that reasons are independent of motivations altogether can account for this in holding that agents must be able to respond to those reasons. However, it is inadequate for different reasons, most prominently because it does not always allow us to ask why some fact is a reason for us to do something (see above III.3.c.). We have seen that these are the two major ways objectivist objections to internalism can go (above II.5.b.). Therefore, at least here, no other account of reasons can be given that shows how (P1') together with (P2) comes out as true under it. We may thus conclude (against the objectivist).

(I): A has a reason to X only if she could arrive at the conclusion to X by soundly deliberating from her actual motivations.

We will see later how different objections to internalism may attack it from a somewhat different angle. For now, we may go on to draw the consequences from this discussion for our theory of rational persuasion.

5. An Internalist Conception of Epistemic Reason-Giving

Remember that we have already structured what a theory of rational persuasion should look like along two desiderata. The first was that it should tell us why rational persuasion is good, whereas force is, in principle, bad. Or, alternatively, it should tell us that this difference is not as deep or clear as it seems to be (see above II.1.a.). The second was that it should tell us precisely what is involved in epistemically and robustly giving reasons and thereby show us what distinguishes them. Or, else, it should show us that the distinction is really not as deep as it seems to be (see above II.1.b.).

I will structure the discussion of an internalist conception of rational persuasion along these two desiderata. First, I will try to work out a precise account of epistemic reason-giving. Second, I will attempt to give an account of why epistemic reason-giving is, in principle, good, contrary to forcing someone to do something. Both of these accounts will proceed on the assumption that internalism is true. In both discussions, I will add minor digressions on which alternative accounts of these phenomena are excluded by internalism. I will conclude the chapter with a

section about the critical significance of internalism and about how it makes us see how distinctions which appear to be sharp actually are not.

Robust reason-giving will have no place in the discussion of this chapter. This is because we have not dealt with the objections to internalism which seem to focus on a model of robust reason-giving yet (see above II.5.c.). This will only be done in the third part.

a. Epistemic reason-giving

Above (II.1.), we already said something about what it is to *give* someone a reason. In epistemic reason-giving, it is to *indicate* that there is a reason for the addressee to do something. Presumably, internalism will not help us to elaborate the character of this act any further. But, it will help us with characterising what is given: reasons. As we have seen (see above II.1.c and II.4.), this is a very important part of a theory of rational persuasion.

The conclusion of the last chapter warrants the assumption that what is given in epistemic reason-giving are internal reasons, i.e. facts that come up as speaking in favour of some action in sound deliberation from the agent's actual motivations (see above II.3.a. and III.4.). This provides us with the following account of epistemic reason-giving:

(EI) A speech act is an act of epistemic reason-giving if and only if the speaker S correctly indicates to the addressee A that A has an internal reason to do X.¹³⁶

A rival account of epistemic reason-giving would proceed on a different, objectivist, understanding of reasons. As it seems to be the most influential objectivist conception today, I will use Parfit's concept of a reason for setting it out. Accordingly:

(EO) A speech act is an act of epistemic reason-giving if and only if the speaker S correctly indicates to the addressee A that A has an external reason to do X.

"External reason" in this definition is to be understood according to Parfit's sense of "reason". The concept itself is indefinable. All such reasons are facts about the objects of desires or aims (see above III.3.b.).

¹³⁶ There are many things one could doubt about this definition. For instance, there may be reasons for considering sincere but unsuccessful attempts at indicating to someone that there is reason for them to do X as epistemic reason-giving. Certainly, a full theory of rational persuasion would need to deal with such doubts and, perhaps, adjust the definition accordingly. This is not the purpose of this thesis, however.

b. Advice and teaching

In this section, I want to briefly illustrate (EI). In order to do so, I will propose to divide up the different ways of epistemically giving reason to do something in two larger types: teaching and advice. I will proceed to say what (EI) has to say about those two types. Finally, I will consider some objections the objectivist may raise.

Unquestionably, there are many different types of epistemic reason-giving. Warning, criticising, and advising each other, to name just a few, can all be analysed as indicating to each other that there is a reason to do something. The distinction between advice and teaching, which I am about to introduce is supposed to divide them up into two larger types so that they can be handled more easily.¹³⁷

Consider the distinction between strongly agent-relative and non-strongly agent-relative aspects of a situation again. Strongly agent-relative aspects are those that change if someone other than the agent imagines herself to be “in the agent’s place” (see above III.3.b.). The epistemic reason-giver accordingly has two options when assessing whether her addressee has reasons to act one way rather than another in her situation. She may think of the situation as it is for the addressee, i.e. including the aspects that are strongly relative to the addressee. Or, she may think of it disregarding the strongly agent-relative aspects of the situation, i.e. as it would be for anyone. This distinction can be put to use in defining two ways of epistemically giving someone a reason. If the reasons that are given only come up in the first perspective, i.e. when the reason-giver imagines the situation as it is for the agent, I propose to speak about “advice”. If the reasons that are given come up also in the second perspective, i.e. when the reason-giver assesses the situation as it is for anyone, we will speak about “teaching”.¹³⁸

The following examples are the paradigms for these concepts. Advice ought to capture cases such as speaking with someone about which of two job opportunities she should take. Absent moral considerations perhaps, whether some feature of a job is a reason to take it obviously

¹³⁷ This implies that there are two concepts of advice. One as a species that is different from warning, criticising, etc. and one that comprises the species of advice and other species of epistemic reason-giving as their genus.

¹³⁸ I am aware that this is not the perfect label for what I want to say. Teaching rather seems to be associated with giving someone reasons for belief. This is not such a big problem though as for objectivists practical and theoretical rationality are not so far apart (see above III.3.c.). “Instruction” would be an alternative but we mostly use it in contexts where it is a background assumption that the addressee has some motivations (consider the instructions for a new coffee maker). “Preaching” may also capture the idea but it has a negative connotation which is unwanted at this point and it implies that the speaker is very insistent on the addressee to accept what she says. This is not implied by giving someone non-strongly agent-relative reasons. I will therefore stick with teaching. For the juxtaposition of “instruction” and “education”, along with a criticism of understanding education on the model of instruction, cf. Lefort 1979, espec. 224.

depends on what skills one has, what one likes to do, what kind of work environment one wishes to have etc. Such reasons are strongly agent-relative. If your partner in communication is giving you such reasons, she is giving you advice.

The paradigmatic example of teaching is a parent teaching her child what not to do. For instance, a father may say to his daughter about to throw small stones off the balcony: “Don’t do that! It could hurt someone.” It seems that the father is not, and need not be, making any efforts to consider the situation as it is for his daughter. He is saying something about the situation that seems to be independent of the person considering it: that throwing stones off the balcony is dangerous and that this is a reason for not doing it.

All other types of epistemic reason-giving may be considered to be either advice or teaching. It could also be the case that one type of reason-giving may be advice in one situation and teaching in another. You can warn someone of a danger that is only a danger to the addressee or of a danger that is a danger to everyone and thus presents a reason to escape the situation irrespective of who is considering the situation. Or, so it seems.

Indeed, it follows from (EI) that teaching is not an act of epistemic reason-giving. According to the internalist, all reasons are part of the strongly agent-relative aspects of a given situation (see above III.3.b.). The normative significance of the facts that are reasons for some action in a given situation derives from their ability to come up in sound deliberation from the agent’s existing motivations. A speaker S can therefore only epistemically give a reason if she indicates that some fact is a reason in that sense. In order to do this, she must refer to an aspect of the situation that may change if someone other than the addressee were considering the situation. On this interpretation, advice, in the large sense, is therefore identical with epistemic reason-giving itself.

This is no doubt an interesting analysis, at least in that it says that teaching is, despite appearances to the contrary, not an act of reason-giving. Because it is interesting in this sense, it should be expected that there are arguments against it. I will review two such arguments in what follows.

The first argument is that the internalist’s conception of advice has unacceptable implications. Judith Jarvis Thomson (2001, 48-50) proposes an example that entertains some resemblance to Parfit’s agony argument.¹³⁹ Suppose that someone is seriously ill and will not survive for long

¹³⁹ She also says, implicitly concurring with Parfit, that what a person ought to do is “objective” rather than “subjective” (Thomson 2001, 50).

unless he takes a medicine that will cause him some pain in the near future. Presumably, it would be good for him to take it but, the problem is, he really does not want to and he does not believe that it would be good for him to take it either – in the terms introduced above: he has no motivation to take it. Thomson holds that it would still be wildly implausible to suggest that we could not advise him to take the medicine. We are concerned with what is good for him when we do that and not with what he wants or what he thinks is good for him.¹⁴⁰

There is a twofold reply that is open to the internalist. First, as Williams, and, following him, Kate Manne have argued, such cases allow for the opposite interpretation as well. They ask us to imagine being in such an advising conversation. What if the person you talk to just says “I don’t care. Don’t you understand? I really do not care.” (Manne 2014, 102; Williams 1995a, 39)? Although there is a sense in which you may go on talking to them, saying that not taking the medicine is just really imprudent or something of the sort,¹⁴¹ it seems that the practice of giving advice has come to an end. If the agent is not in any way motivated to give weight to the considerations you bring forth, trying to advise them really has no point. And, if it could not even have a point unless the agent were motivated to give the advice some weight, this motivation must be presupposed by any serious act of giving advice (cf. Manne 2014, 103; 107-109).¹⁴²

Second, the internalist is in a position to explain the intuitive appeal of Thomson’s argument. This appeal, Manne and Williams may argue, is similar to the intuitive appeal of the agony argument. When we imagine someone not wanting to take the medicine, we imagine them at first not wanting to do it. We still think, and in such a case we should not give that thought up too quickly, that there is some kind of thing we can say to make him understand what not taking the medicine would really mean and that he will come to see that he does have decisive reason to take it. If, on the other hand, he convincingly tells us that he would rather not go on living

¹⁴⁰ Not only with that, of course. In giving advice we should also be concerned with whether the advised action would be unjust or “miserly” (cf. Thomson 2001, 67ff.). In order not to unnecessarily complicate matters, I will stick to cases where what is at issue is simply what is good for the agent.

¹⁴¹ Cf. Williams 1995a, 39. Cf. Scanlon 1998, 371f. for criticism of internalism claiming that saying these things also implies giving reasons. For replies see Manne 2014, 111f. and Williams 2001, 95f. – they basically insist that reasons must have something to do with the explanation of action and rational persuasion. Manne (2014, 96f.; 110f.) speaks straight of switching from an interpersonal to an objective stance when we say such things.

¹⁴² For differences between Manne’s view and Williams’s, see Manne 2014, 104-107. Unfortunately, I cannot go into these subtleties here. For the close connection Williams draws between internalism and advice, see also Williams 2001, 92. This should be separated from discussions that introduce an “ideal advisor model” for the internal reason thesis (cf. Pettit/Smith 2006, 147f.; Tubert 2016, 183-185). These discussions do not concern actual advisor situations but counterfactual, ideal, situations. With Manne (2014, 97, n. 20), in discussing actual acts of rational persuasion, I understand advice to be the real practice of talking to others about what they should do. I will not discuss issues connected with the discussions concerning the “ideal advisor model” in this thesis.

than take that pill, we might not stop trying to talk him out of it but this will not be on the mode of advice but on some other kind of persuasion.¹⁴³

These arguments do not show that Thomson's objectivist analysis of the case is unable to give a convincing account of it. To the first point, for instance, Thomson might reply that the internalist confuses what it is to give advice and what it is to successfully give advice. The latter notion incorporates the idea that the speech act of giving advice has the consequences that it usually wants to bring about, i.e. that the addressee is persuaded.¹⁴⁴ And, it is only this that supposes that the addressee is motivated to give weight to the considerations brought forth.

But, this is no argument that the notion of advice *has to* be interpreted in this way. Manne and Williams are holding that the motivations to give weight to the considerations brought forth is a condition for advising to make sense. That these motivations are, at least possibly,¹⁴⁵ present is a condition on advising the other just as the possibility of an answer is a condition for asking a question to make sense.¹⁴⁶

I see no way to decide this issue here. But, deciding it is not necessary. If it seems that both interpretations are defensible, we may conclude that Thomson has no sufficient arguments against the internalist interpretation of advice and this is all the internalist needs at this point. The arguments for accepting (EI) have already been proposed in the last chapter.

Regarding teaching, one may argue that the internalist conception does not provide an account that satisfies the intuitive evidence. Any account of reason-giving should allow that the father saying to his daughter that she should stop throwing stones suggests a reasonable course of action, whatever the motivational set of the child looks like at this moment.¹⁴⁷ According to (EI), however, such statements do not count as advice if the child has no motivation to listen to her parent and, consequently, they do not count as epistemic reason-giving. The objectivist, on the other hand, has a good account of such teaching speech acts which distinguishes them from advice and still allows them to be acts of epistemic reason-giving. She may say that, in teaching,

¹⁴³ See also Williams 1985, 47ff. on this kind of case. Of course, it does not follow that we would advise this man not to take the medicine if it turns out that he really does not want to. But this is just because we don't want to encourage people to do things we don't think are good things to do by advising them to do so.

¹⁴⁴ Thomson may be understood to suggest this reply in Thomson 2001, 50.

¹⁴⁵ Cf. below III.5.d. on the proleptic function of blame. There may be a proleptic notion of advising too.

¹⁴⁶ This latter point has been an implicit premise in the first of the three argument against Parfit in III.3.c.

¹⁴⁷ For an externalist position that seems to take precisely this kind of situation as its inspiration see Wong 2006, espec. 547-549. This also seems to be the main point of arguments aiming to show that internalism is false beginning with the premise that statements saying that some course of action is morally required give the addressee a reason to follow that course of action whatever her motivations are (cf. Finlay/M. Schroeder 2017, 3.1.1).

agent-neutral reasons are given and in advice agent-relative reasons are given.¹⁴⁸ This explains why in the latter case only strongly agent-relative aspects of the situation count and in the former we may have regard to other aspects of the situation as well.

In fact, the internalist may distinguish (at least) two different meanings of the father's statement that his daughter should stop throwing stones. One is that it may be an "optimistic internal reason statement" (Williams 1995a, 40). That is, not knowing for sure whether his daughter actually has motivations in the light of which some fact will count in favour of some course of action in sound deliberation, the father just assumes that this is the case because it is quite likely, or, we will come to this possibility later (III.5.d.), because assuming that it is the case may actually make his daughter acquire these motivations. In this case, it may be a limiting case of epistemic reason-giving (see below III.5.e.).

Alternatively, such a statement may really not say anything about anyone's reasons at all. Instead, it may just say something about what kind of actions count as a good instance of action type T. Admittedly, this does not square so well with what the father is saying. But, we can imagine him say that good children don't do dangerous things like this. This statement is then of the kind of statements about what a good chess player would do (see above II.3.b.). It says nothing, at least not directly, about whether the addressee has a reason to attempt to be a good child or a good chess player.¹⁴⁹

This is a full account of what speech acts classified as teaching above may mean and how they connect to what is good or reasonable.¹⁵⁰ This is what the argument against (EI) suggested that the internalist could not provide. The internalist account of epistemic reason-giving therefore holds against the second objection too.

Given the results of the last chapter, we may now conclude that (EI) and its implications are correct. Epistemic reason-giving is thus equivalent to giving the addressee internal reasons for

¹⁴⁸ See above III.3.b. on the difference between agent-neutral and agent-relative reasons and its possible relation to the distinction between strongly agent-relative and non-strongly agent-relative aspects of a situation.

¹⁴⁹ Telling someone that this is what a good child would do may still be misleading. As opposed to a good chess player, it is not too clear what a good child is (cf. Williams 1972, 55-62 for the concept of a "good man").

¹⁵⁰ Also, the internalist herself may use the distinction between agent-neutral and agent-relative reasons to distinguish between teaching and advice. On an internalist account too, the full statement of the reason for not throwing stones off the balcony includes no reference to the agent who throws them down. It is only the explanation for why it is a reason for someone not to do this that does contain such a reference. The internalist's concept for this would be different from what we called "teaching" at the outset though, as this kind of teaching does contain a reference to strongly agent-relative aspects of the situation.

action. Consequently, it is also equivalent to giving advice in the large sense. Teaching is no act of epistemic reason-giving

c. Freedom and theory

Let us now turn to the first desideratum for our theory of rational persuasion. We need an account of the trait that makes epistemic reason-giving in principle a good thing, and which will therefore allow us to distinguish it from force, which we view principally as a bad thing. Of course, it may, and it will, turn out that there are cases that are not so clearly relegated to one side or to the other.

I propose that we think of epistemic reason-giving as a good way to engage with others because it furthers their practical freedom. Properly spelled out, this account has two components. One is that furthering the practical freedom of others is good. The second is that epistemic reason-giving is furthering the practical freedom of others.

Practical freedom may be thought of as the idea of people being free in the sense that they are able to take actions that will effectively further some projects, causes, or simple desires they endorse, including the project to reflectively engage one's own projects.¹⁵¹ Although it has largely been presumed that freedom in this sense is valuable, even liberal theorists have recognised the need to give reasons for its importance (cf. Raz 1986, 6-14; Dworkin 2011, 367). Unsurprisingly, these arguments are too diverse and complicated to be considered here. In what follows, I will therefore assume that there are sound arguments that show that in many cases it is good if someone's practical freedom is furthered.¹⁵² I don't think that many will disagree with me on this.

The question of which actions will effectively further some project, cause, or desire of the agent is what we have called a practical question or problem above (II.3.a.). Assistance in providing a good answer or solution to those problems will therefore be, at least in some sense, assistance in realising one's projects, causes, or other aims. Epistemic reason-giving is nothing but providing assistance in the solution of practical problems. It cites facts that count in favour of one

¹⁵¹ This is a rather trivial "positive" concept of liberty: cf. Christman's (2022, 1) introduction of positive liberty as "[the presence of] the capability to carry out one's wishes, or the capacity to reflectively accept one's values, or the recognition by others of one's status as a social being". I don't think it is important that it is a positive concept in that the ability to further one's projects implies the "negative" aspect of being free from interference as well. The aim here is simply to render a common-sense idea of what freedom roughly means. On "positive" and "negative" liberty see Berlin 1969.

¹⁵² I believe that arguments that derive this from an idea of "ethical independence" (Dworkin 2011, 211-213; 368-371) pick up something right.

course of action which may contribute to resolving those problems. Consequently, it contributes (or aims to contribute, at least) to enhancing the practical freedom of its addressee. It consists in helping people to freely develop themselves and what they think is valuable. This is why it is good.

It will be useful to clarify the account further in saying what kind of alternative account it rejects. This alternative account is based on (OE), and, hence, it is “objectivist”.

According to such an alternative account, the distinctive mark that makes rational persuasion good is that it aims at telling others the truth. As we have seen (III.5.a.), on the objectivist account of epistemic reason-giving, (OE), whether some speech act counts as (epistemic) reason-giving depends on whether the fact that is indicated to be a reason for the addressee to do something actually is a reason for that. Whether this is so is decided by the best substantive normative theories.¹⁵³ Given that telling the truth about some matter is generally good (and it must be even more so when the matter is whether something is good) reason-giving is good when and insofar as it tells the addressee the truth about what reasons she has for carrying out different actions in her situation.

It is this account of the worth of epistemic reason-giving that is rejected by internalism. This is reflected in that one of the core tenets of Williams’s philosophy was a profound skepticism of ethical theorising (cf. Moore 2006, 225f.). I will briefly follow this through.

Ethical theories strive to make our ethical judgments more coherent by providing “a theoretical account of what ethical thought and practice are” and “a general test for the correctness of basic ethical beliefs or principles” (Williams 1985, 80).¹⁵⁴ This is most easily seen in Utilitarianism which proposes one large principle, some understanding of maximization of welfare, and methods of calculation which serve to decide in each particular case what ought to be done.¹⁵⁵

In providing tests for correctness, ethical theories tell us what to do. The question Williams asks is: “How can any ethical theory have the authority to do that?” (Williams 1985, 110). One answer that might be given is that it allows us to represent our values in a systematized way and that this will help us resolve conflicting judgments regarding particular cases. Williams calls

¹⁵³ This is implied in Parfit’s (2011a, 43-47) introduction of two “theories” of normative reasons. See also his expression of the hope that “philosophers, or other wise people” would tell us what we have reason to care about (Parfit 2011a, 107).

¹⁵⁴ This is what Williams understands by a “positive” ethical theory. For “negative” ethical theories see Williams 1985, 82.

¹⁵⁵ For Williams’s critique of utilitarianism see Williams 1972, 82ff.; Williams 1973; Williams 1985, 116ff.

this “rationalistic” (Williams 1985, 111). This targets a conception of ethical theorising that is inspired by Kantian contractualism whose aspiration Williams summarises as laying bare the foundations of our ethical judgments (cf. Williams 1985, 111f.). I will not speak about this here because this sort of theorising would still proceed on an internalist account of reasons. Another reason to accept theory’s authority might be that, in its theoretical mode, it detaches us from the prejudices of ordinary life and allows us to view our lives from the outside, objectively (Williams 1985, 123). This latter feature may be a reason to think that the ethical theory may be more likely to tell us what the truth about what we should do is. Basically, this is the objectivist account we just introduced.

Williams’s main argument against this account is that it rests on a misunderstanding of what it is to be a practically rational being. A rational agent should indeed reflect on what she should do and thus look for reasons which speak in favour of one course of action. But, this is not always to look for a further reason that justifies, or, perhaps better, sustains the first reason’s force, which is, in a systematic way, what ethical theories do. Williams expresses this in writing: “[t]heory and prejudice are not the only possibilities for an intelligent agent” (Williams 1985, 124). Often, it just seems to be not only redundant but wrong to look for further justifications for some actions than the ones that we intuitively have: “‘You can’t kill that, it’s a child’ is more convincing as a reason than any reason which might be advanced for its being a reason.” (Williams 1979b, 81; cf. Moore 2006, 236f.) The problem with an account based on (OE) that views the paradigm of good epistemic reason-giving in ethical theories is that it cannot show how such reasons depend on reflection in any way without giving some further reason of why they should be a reason.

This argument parallels the argument from the normative question (above III.3.c.). We can show that we are reflective agents in critically thinking about the normative significance of facts such as that that is a child. If the concept of a reason is specified in the internalist’s sense, we have a good picture how. We may explain it in referring to sound deliberation from our actual motivations and we may then tie it back to a strong moral belief that children should not be killed. The objectivist, at least if she holds, with Parfit, that the concept of a reason is indefinable, must say that there is really nothing to think about here, or else, that there is another reason that is the reason why the fact that that is a child is a reason not to kill it. Both are implausible. Therefore, the objectivist picture, implying that the paradigm of epistemic reason-giving is giving others a systematic theoretical account of why they should do something, should be rejected.

Regarding epistemic reason-giving, we have now dealt with the two desiderata for a theory of rational persuasion. We have specified epistemic reason-giving as indicating to the addressee that she has (internal) reason to do something. It is, in principle, a good thing to epistemically give someone a practical reason to do something because it enhances the practical freedom of the addressee. I want to conclude this chapter in drawing some consequences of this discussion.

d. Bluffing and blaming

First, I want to demonstrate the critical significance of (EI). If a sort of action is recognised to have ethical worth, there is the possibility that actions which do not have this worth are confused with those valuable actions. This may lead us to attribute a kind of force or importance to actions that we would not attribute to them if we understood them right. Showing that our evaluation of some actions rests on such misunderstandings is what I call “critique” here.¹⁵⁶ As epistemic reason-giving is presumed to be, generally, a good thing, it is conceivable that other types of action are confused with it, and, consequently, credited with force they don’t deserve. As we will see, that internalism has the critical potential of showing us where and why this actually happens is one of the features of internalism that Williams was especially keen on emphasising.

It should be noted that the criticism proposed by internalism is itself a type of rational persuasion. It provides us with a better picture both of what our motivations to further the liberty of others and our motivations to accept the assistance of others in resolving our practical problems require of us by making us better understand the nature of what we are really doing in certain situations.

Williams discusses his position in “Internal and external reasons” using the example of Owen Wingrave (cf. Williams 1979a, 106ff.; above II.4.). Owen’s family tells him that he should join the military because this is what family honour requires. According to Williams, this is a paradigmatic case of an external reason statement. Knowing that Owen does not really care about family honour, it purports to recite a reason that is independent of Owen’s caring. Given that the internal reason thesis is true, there is no such independent reason. It is only in having regard to Owen’s motivations that we can say whether family honour is a reason for him to do something, in this case to join the army. Consequently, the family’s recital of the purported reason to join the army, i.e. family honour, is merely a “bluff” (cf. Williams 1979a, 111).

¹⁵⁶ This concept of “critique” is modelled on Williams’s (1985, 112) notion of “transparency”, i.e. that “the working of [the] ethical institutions [of a society] should not depend on members of the community misunderstanding how they work”. Cf. Williams 2002, 219-232 for Williams’s own, slightly different, notion of “critique”. For a classical discussion of how theory can be critical see Horkheimer 1937.

To get a better grip on this notion, let's consider the most familiar meaning of the term "bluff": it is a strategy in poker. One description of what happens in poker when someone bluffs is that she pretends to have high cards in order to win the round. On this level of description it is equivalent to a lie directed at making someone else do something which will help one achieve one's own goal. But, there is another way to describe what is happening which helps us distinguish bluffing from simply lying. After all, the poker player is not pretending to have particular high cards in her hands. What she is pretending is that her cards are such that she will win this round. Now, if she is believed and her opponent acts on this belief, she will not only achieve her goal but what she pretended to be true will be true in fact. The success of the pretense stops making it a pretense.¹⁵⁷

If we take Williams's use of the notion of a bluff seriously, we can see that this sort of structure is exactly what might happen with Owen. If he believes his family, he will think that he has a reason to join the army: that this is what family honour requires. Either this belief itself or some motivating state of mind it generates might motivate Owen to join the army.¹⁵⁸ These are the motivating states of mind required for an internal reason to act to come about. Consequently, if successful, the pretense of an external reason statement makes it stop being a pretense because it is now a true internal reason statement.

It may be asked why Owen should believe his family saying that he has a reason to go to the army. This is where the content of their pretense comes in. The external reason statement has the appearance of being an act of epistemic reason-giving. Epistemic reason-giving is, in principle, a good thing because it enhances the practical freedom of the addressee. Owen, and his family, however, may rather understand the act of reason-giving as an objectivist would understand it and think that a reason statement just states that a fact is a reason for doing something without being connected to the addressee's motivations. Thus when stating that Owen has a reason to join the army, his family may present this as a true normative fact that is unconnected to Owen's motivations. And, we often do believe statements that pretend to utter something that is true. That they have a wrong understanding of what reasons are might therefore both explain why Owen's family feels justified in telling Owen to join the army and why Owen thinks that he ought to believe them.¹⁵⁹

¹⁵⁷ Cf. Lindeman's (2020) analysis of the speech act of conservatives.

¹⁵⁸ For the possibility of motivating beliefs see above II.3.a. and Williams 1979a, 107.

¹⁵⁹ A debate in the vicinity of this is the moral testimony debate (see Hlobil 2020, 987-989 for a recent summary). Since this debate is concerned with morality and not with reasons, it does not clearly concern the same issues with which we are dealing here. I will therefore disregard it.

External reason statements of this type therefore derive their force from appearing to be reason-giving and, hence, in some sense, a good thing, when really they are not. Internalism can show us that their looking like rational persuasion is merely an appearance. This is its critical force.

Bluffs are not the only speech acts that are singled out this way. Perhaps even more famously, Williams criticises the practise of blaming one another as involving something other than it pretends to involve. I will retrace this here too.

Normally, blaming is something we do to each other when we are in “deliberative communities” (Williams 1985, 215).¹⁶⁰ This means that we have some sort of practical issue and try to resolve it together. If someone does something that inhibits the resolution of this problem, it will be appropriate to blame them. A paradigm case might be the behaviour of a sports team (cf. Williams 1995a, 40). Suppose a football player shoots, and misses, from a mediocre position rather than passing the ball to her team member who is in a position in which she will most likely score. The other team members will have the standing to blame her on the grounds that she did something that was a hindrance to the resolution of their common problem: winning the game. Such a “deliberative community” presupposes a community of interest (interest taken in a large sense).¹⁶¹ The members of the other team could not blame our soccer player because they are not part of this community. This is so because their problem is a very different one: how *they* could win the game.

Contrary to this, we do seem to blame others when we do not have such a community of interest. Williams invites us to consider an example: a man treats his wife horribly. When asked about this behaviour, he sees no reason to stop.¹⁶² He is just not motivated to treat his wife better. Williams supposes that we might still be tempted to blame him for this – even though he is not on our team, i.e., perhaps, the team of people who want to make sure that everyone is treated with the respect they deserve.¹⁶³ If the blame this man receives is anything like the blame the

¹⁶⁰ Likewise, Darwall (2017, 270) takes blaming to happen from the standpoint of the “moral community”. What exactly makes the moral community the kind of community it is, is a different, and difficult, question.

¹⁶¹ I do not mean to presuppose any position in debates about collective intentionality here (see Schweikard/Schmid 2020). That this presupposes a community of interest follows from the requirement that the individuals thinking about what to do are attending to the same practical problem, internalism about reasons, and the view that practical reasoning is considering reasons for or against some course of action (for the latter see above II.2.d.). Whether thinking about the same problem is what jointly deliberating is, is a further question that is not answered here.

¹⁶² As Manne (2014, 103, n. 37) correctly points out, it is well worth reflecting on that, in Williams’s discussion, it does not seem as though the wife would enter the conversation. I do think that we should urgently take this matter into account, as it promotes thinking that the victim could not play a role in the rational persuasion of the offender against some value. I hope that the discussion in the last part will allow us to see what her role might be in rational persuasion. I will not be able to draw this conclusion explicitly though.

¹⁶³ Or the team of good spouses, see Manne 2014, 109. For Darwall, this team might be the moral community.

soccer player receives, it cannot be appropriate. So we need a different account of what the point of blaming the man is.

Williams delivers such an account in saying that blame functions “proleptically” here (Williams 1995a, 41f.). We treat the man as if he were already a member of our deliberative community and blame him for not contributing to the problem we all have. We do this in the hope that, if we treat him like this, he might actually come to participate in the community. He might believe that he belongs and that he has reason to contribute to the solution of this problem (cf. also Williams 1985, 214f.).

The structure of this practice is isomorphic to the structure of the bluff. We act as if something were true that is not true regarding this man (that he has reason to solve the same problems that we all want to solve) in the hope that he will believe us, that he will then, in virtue of this very belief, be motivated to solve this problem with us and that, in virtue of this, what we pretended in the beginning will have become true. Here too, we may do this in pretending that the reasons someone has are independent of what they want. The same critique that applies to bluffing applies to blaming too. We have thus singled out two practices whose functioning presupposes that their participants (or, at least the addressees) believe that they are a kind of epistemic reason-giving but concerning which closer examination shows that they are not.

e. Between force and persuasion

The example of the man treating his wife badly has shown that not only do we sometimes tell others that they have a reason to do something or blame them for not having done something without being entirely sure whether they have, or would have had, internal reason to do it. In certain cases, we do not think that this is a bad thing. After all, we could not just let the man get away with his awfulness. Consequently, perhaps contrary to the case of Owen and his family, the internalist critique of blaming will not make us give up this practice.

This shows that there is no sharp boundary between rational persuasion in the form of epistemic reason-giving and trying to force, or manipulate, people into doing something. It seems that there are speech acts, such as blaming someone for immoral behaviour despite motivations to the contrary, that are rather on the margins of both force and persuasion. They are neither really cases of rational argument nor do they clearly belong to the side of (bad) force or deception. Williams regarded this as a good thing. He strongly emphasised internalism’s ability to show

us how the apparently sharp distinction between mere force and rational persuasion may be challenged (Williams 2001, 96).¹⁶⁴

We can now conclude this part by summarising its main points. *First*, objectivist objections to internalism fail. *Second*, epistemic reason-giving takes the form of advice in a large sense. *Third*, what makes rational persuasion in the form of epistemic reason-giving good is that it enhances the practical freedom of its addressee. *Fourth*, internalism both shows us how these points can be put to critical use and that the distinctions between rational persuasion and its opposite – force – is not as clear as it may initially seem.

¹⁶⁴ A great emphasis on this distinction and other sharp distinctions is taken to lie at the basis of “morality system” and is given the title “purity” by Williams 1985, 215f.

IV. Externalism and Robust Reason-Giving

This part will consist in an argument to the effect that internalism about reasons is not correct after all. The argument has two parts. I will begin by re-examining what we have characterised as robust reason-giving in part one. There seems to be a problem of how to give an adequate account for it and I will suggest that neither the attempt to reduce it to triggering reason-giving nor internalist explications are satisfactory. I will conclude that we should be open to the idea of external reasons for giving such an account. We should expect this account of external reasons to be intersubjectivist (see above II.5.c.), i.e. it should explain reasons in terms of the motivations of other people.

The second part of the argument will consist in suggesting that there is an account of external reasons that is defensible. I will consider two intersubjectivist objections to internalism. Both fail but there are lessons to take away from their failure. Consequently, I will propose my own account of external reasons which, I think, can meet Williams's challenge. I will conclude this part with a small chapter on the consequences of my account of external reasons for a theory of rational persuasion.

1. The Problem with Robust Reason-Giving

Recall our characterisation of robust reason-giving (see above II.1.b.). It consists not in indicating to the addressee that there is a reason for her to do something but rather in the creation of a new reason. It seems, we have said, that the act of robust reason-giving *is* the reason that is given. Examples are A demanding B to get off her foot, or my asking someone to read this thesis.

This is to some extent puzzling. As Enoch (2011, 1; 2014, 300) remarks, there seems to be some kind of normative magic going on here in the precise sense of magic being that only saying words can make a difference to how things are. If I am successful, the normative situation of the ones I ask to read this thesis is substantively altered through my request. They now have a reason to read it. This is puzzling in two ways. First, it is unclear how someone could acquire that power for normative magic, i.e. what has been called the "standing" to request or demand something.¹⁶⁵ Second, and this is the question with which we will be dealing here, it is unclear *what it is* for someone to have that power. We need to clarify what these acts are which seem

¹⁶⁵ See Gilbert 2018 for a treatment of this problem regarding demands.

to involve magic in order to see how and why they do seem to involve such magic. I will call this the “problem of robust reason-giving”.

There are, I believe, three broad ways of dealing with this problem. The first is to defuse it in showing that what we have called robust reason-giving does not involve anything that is akin to normative magic at all. It only *seems* to create a new reason when really it depends on an already existing one. This is the route taken by David Enoch.

The second is to accept that in robust reason-giving a reason is created and to explain the creation of that reason in internalist terms. Consequently, such a reason would be explained in the same broad way as any other reason. The normative magic will not be defused, strictly speaking, but it is no more mysterious than the force of any other reason.

The third is to take the appearance of normative magic seriously. The reason created in robust reason-giving would therefore be neither reducible to another reason, nor would its normative force be susceptible to explanation on the model that applies to other reasons. As the explanation of the force of reasons created in robust reason-giving would then differ from the internalist’s explanation of other reasons, this strategy must embrace some kind of externalism.¹⁶⁶

I will go through these three approaches one by one. I will try to argue that neither the first nor the second could give us a satisfactory account of robust reason-giving. Therefore, we should put our hope in the third strategy.

2. A Reductive Solution

David Enoch proposes to think of robust reason-giving as a species of triggering reason-giving. It is therefore of the same basic kind as the neighbourhood grocer raising the price of milk (for the example see above II.1.c.). There is already a reason for the agent to do something, X, if something else, Y, happens. Demanding, or otherwise robustly giving someone reason, is then the event Y that fulfills the conditional implicit in that reason and therefore triggers A’s reason to do X (cf. Enoch 2011, 9f.). A request to read my thesis would then only be what makes a conditional true that is inherent in a reason you already have, e.g. the fact that we know one another may be a reason for you to do some things if I request them.

¹⁶⁶ There is a fourth strategy that I will treat as a subcategory of approaches falling under the third strategy. It may be claimed that really the explanation of reasons created in robust reason-giving is the model for all explanations of reasons. Something like this may have been proposed by Emmanuel Lévinas (1961, espec. 241f.: “Dans l’accueil du visage la volonté s’ouvre à la raison.” (emphasis omitted)). I will say some things about why I don’t think this strategy could succeed in passing (see below IV.3.b.) but I will not engage as substantially with it as I will engage with the other three proposals.

For Enoch, this seems to be the way to go. The only alternative would be to accept that demanding, requesting etc. could create reasons for others in the stronger sense that their reason for action would be partly constituted by the request/demand. This is normative magic and highly counterintuitive for Enoch (2011, 11).¹⁶⁷

But, defending this view, Enoch must meet another challenge. Robust reason-giving seems to be very different from both epistemic reason-giving and triggering reason-giving (see above II.1.b. and c.). This impression may be false. But, if it is, there should be an explanation of why it seems as if it were so different. Enoch tries to meet this challenge in giving an account of robust reason-giving that is supposed to show that it is a special case within the category of triggering reason-giving (cf. Enoch 2011, 13f.).

Roughly, the idea is that, in robust reason-giving, it is the intention to give a reason itself that triggers the reason (cf. Enoch 2011, 14-18). More specifically, Enoch proposes the following account (Enoch 2011, 15):

(RR): A attempts to robustly give B a reason to X only if (i) A intends to give B a reason to X and communicates this intention to B; (ii) A intends B to recognise this intention and (iii) A intends B's given reason to X to depend in an appropriate way on B's recognition of A's communicated intention to give B a reason to X.

“Giving reason” in those clauses should be understood, if anything specific needs to be understood thereby, to refer to triggering reason-giving. The first intention is thus to fulfill the antecedent of an if-clause implied in an already existing reason (Enoch 2011, 18). The attempt succeeds only if there actually is a reason that is triggered by A's speech act (Enoch 2011, 16) in the appropriate way (Enoch 2011, 17; this is the “success condition”).

So, the idea is that what my asking someone to read this thesis consists in, leaving aside complications that are not of much interest here, is me intending to give them a reason to read it; and me intending that this “given reason [...] depend in an appropriate way” on their recognition of that intention; and that I thereby actually trigger a reason to read it.

Many objections against this account have recently been brought forth by Ezequiel Monti (2021). There is no place here to deal with each of these objections. I will therefore only discuss the one I consider to be the strongest.

¹⁶⁷ In fact, Enoch (2011, 9-12) provides a little more argument for this point. He opposes a conditional and a constitutive model for a explaining a fact's normativity. I will leave this out because it would significantly complicate the argument and what is most important is rendered in simply speaking about normative magic. James Lewis (2018, 12) makes a similar argument.

According to this objection, Enoch's account cannot distinguish between robust reason-giving and threatening someone (Monti 2021, 3744-3747). That there is such a distinction has often been remarked, specifically in legal theory where this distinction is of obvious importance.¹⁶⁸ Enoch himself (2011, 5f.) considers it to be a condition of adequacy for any account of robust reason-giving that it allows for that distinction.

According to Enoch, threats are different from robust reason-giving in that the reason triggered in threatening someone does not normatively depend on the intention to give the addressee a reason. Rather, it is a prudential reason that depends on the probability of the speaker carrying out the threatened action. It is thus not the intention to give a reason that fulfills the if-clause but something that is indicated by that intention. Roughly spoken, Enoch holds that, just as warnings, threats are an information of upcoming danger (cf. Enoch 2011, 14; 18).

Threats, however, are more than merely informative of an upcoming danger. Even if what triggers a prudential reason of the addressee is, strictly speaking, not the threat but the probability of an upcoming action of the speaker that has bad consequences for the addressee, the threat is not only epistemically connected to that probability, as a warning of a danger would be. The threat itself, or, more precisely, the intention to threaten the addressee plus the intention that the addressee recognise that intention and lets it play an appropriate role in her practical reasoning, may change the motivational psychology of the speaker such that the probability for the hurtful action arises in the first place (cf. Monti 2021, 3745f.). Often, the one issuing a threat would not do what they threatened to do if they wouldn't have issued the threat. The triggering of the addressee's prudential reason therefore indirectly, through this further causal relation between the threat and the probability of hurtful action,¹⁶⁹ depends on the speaker's threat (including the relevant intentions). This seems to fulfill the success condition too. A threat may actually trigger a reason for the addressee, though only indirectly (Monti 2021, 3745). Therefore, Enoch would be forced to admit that paradigm cases of threats are instances of robust reason-giving. This contradicts the premise that there should be a sharp distinction between threats and robust reason-giving.

Enoch may reply that such indirect triggering should be ruled out. It needs to be the speech act itself that triggers the reason and not anything that causally depends on it (cf. Monti 2021, 3748). This would rule out threats but not my request to some people to read this thesis. In this

¹⁶⁸ A locus classicus is Hart 1961, 19f.

¹⁶⁹ On counterfactual theories of causality, which I have implicitly used here (see the seminal piece D. Lewis 1973). I don't think anything depends on the choice of theory of causality here.

case, it is the request itself that triggers the reason to read the thesis and, hopefully, not the probability that I get mad at those people if they don't read it or something of that sort.¹⁷⁰

This is not enough, however, to rule out all counterexamples (cf. Monti 2021, 3748 for the following). Consider the following case. Susy wants to break up with Derek and plans to do so during dinner. Derek suspects that this will happen and he wants to stay with Susy. When Susy tells him that she has made reservations at a restaurant for the next night, he correctly thinks that she will break up with him there. He goes on to say "I love you", which they have not said to one another yet. He intends this to make their relationship more significant or important, which will make actions with regards to it, such as breaking up, more significant too. In this, he intends to trigger reasons Susy has that require her to consider more significant actions more strongly, which, in this case, may mean postponing the break-up. He will intend Susy to recognise the intentions to make the relationship more significant and he will intend to thereby give her reasons to consider actions related to it more strongly and he will intend these intentions and what they constitute to play a part in her reasoning.¹⁷¹ Lastly, having a more significant relationship may actually be a reason to consider actions related to it more strongly.

This case fulfills all the conditions set out by Enoch and involves no indirect triggering through causal relations. There is mediation through constitutive relations though. What triggers Susy's reason to postpone the break-up is not the intention to make her do that but the increase in the significance of their relationship which is partly constituted by the intentions to increase its significance. Derek's speech act is therefore an act of robust reason-giving through mediation of constitutive relations. Such mediation is not ruled out by Enoch's account and it hardly could be. Surely, many acts of robust reason-giving involve such relations. It is an important part of the order of a police officer to stop your car that it will constitute an illegal action if you keep on driving.

There is therefore nothing to exclude such cases from the class of robust reason-giving on Enoch's account. Clearly, Derek's statement is not an act of robust reason-giving as we set it out though. Indeed, it seems to be rather on the side of threats than on the side of demanding,

¹⁷⁰ It is still dangerous for Enoch to limit his account of robust reason-giving in this way. It may then rule out speech acts we want to be included in our account of it (cf. Monti 2021, 3749). I am not entirely convinced that this is the case, which is why I will set that worry aside here.

¹⁷¹ Derek may or may not intend Susy to recognise the intention to make her postpone her break-up more specifically – this is irrelevant because it is not required by Enoch's conditions if the action for which he intends to give her a reason is only described as considering actions related to their relationship more strongly.

requesting, etc. Enoch's account is therefore unable to deal with major counterexamples and should be rejected.

3. Internalist Answers

In this chapter, I want to see whether internalism can deal with the problem of robust reason-giving. First, I will consider answers to it that are straightforward in the sense that they do not presuppose an elaborate theoretical background. After having presented evidence that such answers are not satisfactory, I will move on to consider one internalist answer that is not straightforward in this sense because it presupposes a Kantian theory of practical reason. Here, too, I will present one reason for thinking that this answer is problematic.

a. Straightforward answers

The internalist may provide a much simpler account of robust reason-giving, compared to Enoch's reductive account, showing how it is special with regards to standard examples of triggering reason-giving. We may just say that what distinguishes robust reason-giving from the grocer-case, for example, is that the speech act in the request case is the "real" reason for acting whereas in the case of the grocer it is not. You read my paper because I requested it, i.e. this is the fact that stands out in your perspective as speaking in favour of you reading my paper. But I eat less bananas because they are more expensive. It is the price of the bananas that stands out to me and not the grocer's raising that price. This sustains the intuitive difference we made between robust and triggering reason-giving. Robust reason-giving is not something magical on this account because there is a straightforward explanation for why my request constitutes a reason for my addressees: there is a sound deliberative route from some of their motivations, e.g. friendly feelings towards me, to the action of reading my paper. All seems to be well then.

But, saying that the speech act just *is* the reason for doing something is not enough to make the difference satisfactorily. Suppose that you just tell me plainly that I suck. I will probably just leave you right there. My reason for leaving will be that you have insulted me. This is what stands out to me (that I don't want to be around people who are mean to me explains why it does but it is not itself the reason, just like my friendly feelings towards someone are not my reason for doing what they request, the request is). But, insulting someone is intuitively no act of robust reason-giving. You don't command, request or demand anything of me.

There is an obvious way to modify the internalist account, however. We may say that, in the case of robust reason-giving, the motivation that is part of the explanation of why it is a reason is in some ways special. The internalist may say, for instance, that what explains the normative

force of order, requests, or demands is some motivation for doing what someone else orders, requests, or demands. Friendly feelings might be such a motivation to do what one's friend requests. In the insult-example the explanation of why the insult is a reason to leave does not refer to such a motivation, however, but to some desire of mine not to be in situations in which I am under pressure, or to my acceptance of some conception of honour.

The problem with this suggestion is almost as obvious though. If we want to understand the motivation to do what someone orders, demands, or requests, we already need an understanding of these speech acts. This is, however, precisely what we are seeking in trying to construct an account of robust reason-giving. This motivation is therefore no less obscure than robust reason-giving itself. The proposal is simply circular.

The motivation in question would need to be construed such that it does not include a reference to the specific reason-giving-quality of acts of robust reason-giving. The question is therefore whether we can analyse demands, requests etc. such that a purely factual part and a part of normative significance can be distinguished.¹⁷² The factual part could then be part of the object of the motivation that explains the normative significance of the act of robust reason-giving. Let's see how this goes.

The obvious candidate for the factual part is a statement of imperative form, or any act that expresses something usually expressed in statements of the imperative form. The idea is that there are things people say that impose conditions of satisfaction on subsequent actions such that some actions will count as doing what is said.¹⁷³ Thus, when I say "Pass me the salt, please", it will be ascertainable after whether you have done what I have told you to do, i.e. whether you have passed me the salt.

The internalist may add that such a statement is a reason to do something for the addressee only if there is a motivation of the addressee such that this statement will come up as speaking for some action in ideal deliberation from it. Such a statement is an act of robust reason-giving, the account may go on, only if the motivation that explains why it is a reason is simply to carry out actions which count as doing what the speaker tells the addressee to do, i.e. which satisfy the conditions of satisfaction implicit in her statements. This would be the internalist's account of robust reason-giving.

¹⁷² This is what Williams (1985, 155-157) denied for "thick ethical concepts".

¹⁷³ For a short introduction of the idea of "conditions of satisfaction" along with "direction of fit" see Searle 2001, 37-39.

One addition to this view is certainly needed. Not all statements that impose conditions of satisfaction on someone's action are requests even where there is the addressee is motivated as required by the internalist. Consider the case of Ancient Roman penal law. It is possible to ascertain whether someone conforms to it or not, yet they certainly do not count as – present – demands that someone do something.¹⁷⁴ And if someone were motivated to do what is required by it, this would still not be the case. The internalist concept of robust reason-giving should be able to account for this fact.

One possibility for doing so would be to combine Enoch's intention-based account with the internalist account. This has been proposed for requests by James Lewis (2018, 12-14). Roughly, his story goes, one's valuing someone discretionarily¹⁷⁵ explains the reason-giving quality of the valued person's speech acts carried out with the intention to give the addressee a reason to do something. One example of such valuing might be the friendly feelings that we have already invoked as a candidate for explaining the normativity of my request for someone to read this thesis. If I have the appropriate intentions to make my addressees have a reason through my request, presuming that they value me to the extent required for them to make it reasonable to do what I request them to do, and they actually value me in this way, my request is a reason for doing what I request on this account. The intuition that Roman penal law no longer counts as demands may be dealt with in saying that there is no longer anyone who has the intention to give anyone else a reason for adhering to it. In the weak sense in which we can still imagine Roman lawgivers and their intentions, we may actually say that they constitute an act of demanding someone to do something – although, we may say, not presently. And, this is certainly alright.

I actually think that something of this sort might be the correct view of robust reason-giving. I have some worries though which will lead me to set out a different, externalist, account. Roughly, the idea is that in solely acting on our friendly feelings or other ways of valuing others, such as respect for someone's achievements or for institutional requirements, we don't take demands or requests seriously as demands coming from the outside. In a way, we are still too self-concerned if we do so.¹⁷⁶

¹⁷⁴ For a theory of law that largely understands it in terms of a broad concept of commands, cf. Kelsen 1960, 4-10.

¹⁷⁵ For this concept see James Lewis (2018, 6-8): roughly, it means that an agent is in a position to choose whether to fulfill this value or whether to do other things without facing reproach.

¹⁷⁶ This seems to be part of Korsgaard's thought in 1996a, 134-136; 140-142.

To illustrate this, consider first two kinds of reasons for which one could fulfill my request to read this thesis. First, one could do so in order to make me indebted to them so that I would return the favour when needed. Even though we surely often do reason like this, in a way, this is not taking the request seriously. Our reason for fulfilling the request is then not the request itself but the probability of the addressee returning the favour. A motivation to further one's own goals, which is what explains the normative significance of this latter fact, might be enough sometimes to explain why it is a reason for someone to do something. It is not enough, however, to explain the normative significance a request, *qua request*, has for its addressee. And it is the explanation of this normativity that we are seeking.¹⁷⁷

Second, you could fulfill the request because you actually like me, or value me in some other way. You would then do it “for me” (cf. J. Lewis 2018, 7) and not “for yourself”. This seems to be both ethically more praiseworthy and truer to the request. You are concerned with me and not simply with your own benefit when you do so. This seems to be a way better candidate for explaining why requests just are reasons for the addressee to do something.

It seems to me, however, that this is not enough. It still lacks sufficient concern with the request as a standard of action coming from the outside. When you only fulfill the request “for me” or “for my sake”, you are still promoting only what you deem valuable – showing respect for me, perhaps, or furthering my well-being. But, if this is your entire motivation for doing so, it still seems as if you are not taking me seriously in a way. A request is not just an opportunity for doing someone a favour, or doing something one thinks is good. It is an expression of what the speaker herself deems to be valuable. If we are to take this seriously, it seems, it is these values, the speaker's, that we should pursue in fulfilling the request, not our own.¹⁷⁸

Another way to put this might be to say that when we are fulfilling the requests or demands of others, we are not only acting *for* them. Sometimes, at least, we are acting *on their behalf*. In a way, our actions in fulfilling the demand are no longer our own. They are the other's (this idea is further discussed in IV.5.b.). In other words, we let ourselves be “guided directly by another person's will” (Hanser 2015, 95). This letting the other take control of ourselves seems to be the most genuine way to regard the demand as creating reasons for us. It, rather than our concern

¹⁷⁷ Markovits (2010, 203) calls the implicit general thought that our actions are only morally worthy when “performed for the right (motivating) reasons” the “Right Reasons Thesis”.

¹⁷⁸ Samuel (2021, 14) may be understood to make a similar point when he emphasises that we should take others seriously as “subjects”. I think, at least part of what he means thereby is that we should treat each other as subjects of motivations or of values and not simply as valuable beings. I believe that this is another way, unrelated to requests, of working out the thought that I am presenting here.

with what is good for the other, our respect for the other's institutional position, or whatever the motivation would be that explains why one should fulfill a demand *for* the one issuing it, is the reason-giving force of the demand or request *as such*.¹⁷⁹

It should be clarified that doing an action “on behalf of the other” need not be inconsistent with doing it for the other. We might, and I think we often do, let the other's projects guide our actions out of our love, respect, etc. for them. But this love is then a reason for letting the other's will guide our actions. It must be distinguished from the reasons that we have in virtue of letting the other's will guide our actions. It is these latter reasons, I contend, that constitute the peculiar normativity of acts of robust reason-giving (cf. Hanser 2015, 114). I will later explain how I think this works exactly (IV.5.c.).

Another point I want to mention right away is that, in putting these last thoughts forward, we have actually introduced a shift in how we think of an act of robust reason-giving as creating reasons for its addressee. If the point of following requests, demands, etc. is to let one be guided by the other's will, it is not so much that the request *is* a reason for the addressee for doing something. That would need an explanation in terms of the addressee's motivations. It is that the will expressed in these requests, demands, etc. is itself a standard according to which we judge whether some fact is a reason for acting or not. It *explains* why some other fact is a reason for action much in the same way that the motivations of the agent do. As Hanser (2015, 107) puts it: “The point is [...] that doing another's bidding involves [...] acting for someone else's reasons instead of one's own.” I will further spell out how one may think of this in the last chapter.

If this is correct, the more sophisticated Lewis-style internalist account for robust reason-giving fails too, as it cannot account for the way in which we respond to what others want or value that seems to be the most genuine way of fulfilling another's demands or requests. As I said, this conclusion is only preliminary. More should be said on this and this is an important way in which this thesis would need to be developed further. Also, I don't doubt that there are more elaborate accounts that are open to the internalist. And, of course, I cannot show that none of them provides a good analysis for acts of robust reason-giving. However, I hope to have shown that it is not completely straightforward that the internalist has a satisfying analysis of robust

¹⁷⁹ Hanser (2015, 96) identifies this as one desideratum, a theory of “doing another's bidding” (i.e., I believe, robust reason-giving in our terminology) has to satisfy.

reason-giving. The challenge to give a satisfying account of robust reason-giving therefore remains to be answered for the internalist.

b. Kantian approaches

The challenge of robust reason-giving has been prominently considered by at least one theorist inspired by Kant: Stephen Darwall.¹⁸⁰ In this section, I want to see whether he can give us an adequate explanation of robust reason-giving. If this is so, an internalist account of these phenomena is possible as Darwall is an internalist (cf. Darwall 2017).¹⁸¹ Of course, I can only render Darwall's most basic ideas. This section is not supposed to provide a full criticism of his view but rather only the attempt to set out the main points a full criticism would need to address.

The most interesting concept Darwall employs for his theory is the concept of "second personal reasons". Remember the toe stepping case (above II.1.b.). When, instead of merely indicating to me that I am standing on your foot, you demand me to get off, we said that the demand *is* the reason you give me to get off. This is not only what we characterised as robust reason-giving. The reason given also makes a vital reference to *you*, i.e. the addresser or the second-person. As Darwall (2006, 8) puts it, "They [i.e. second personal reasons] simply wouldn't exist but for their role in second personal address."

Not all demands generate reasons simply like that. Rather, they presuppose that you have the "standing" or the "authority" to demand something like that. In other words your demand must be legitimate. I could not just demand my friend to sell her house in order to finance my, admittedly not too promising, tennis career. Similarly, a recruit may not demand that an officer salute him. He does not have the authority to do that (cf. Darwall 2013, 137).

This relation of having the authority to demand something from someone and thereby giving them second-personal reasons to do something implies the accountability of the agent to the one issuing the demand. If a soldier doesn't listen to her commander giving her orders, she "disrespects" her authority – she just acts as if it were not there. And, if the foot-stepper does not answer to your demand to get off, she does likewise. Answering to the authority about acting in accordance, or not in accordance, with the second-personal reason given, is thus a form of respecting this authority, i.e. of acknowledging that it is there or of recognising it (cf. Darwall 2006, 60).

¹⁸⁰ Christine Korsgaard (1996a, 132-145) also addresses topics in the proximity of this problematic. I focus on Darwall here because he addresses the speech acts we have classified as robust reason-giving more explicitly.

¹⁸¹ On the relation between Kantianism and internalism see above II.5.a.

So, we have come to consider four concepts: second-personal reasons, demands, authority and accountability. Darwall (2006, 12; 2013, 141) claims that these are interdefinable and irreducible. This means that they are all defined by one another and cannot be defined in any other way. Thus, the definition of second-personal reasons that Darwall (2006, 8) gives us is as follows: “A second-personal reason is one whose validity depends on presupposed authority and accountability relations between persons and, therefore, on the possibility of the reason’s being addressed person-to-person.”¹⁸² The idea is that these concepts explain an entire stance – the “second person standpoint” (Darwall 2006) – or a way of relating to one another. A way of relating in which a number of attitudes makes sense that we address to one another, such as resentment and gratitude, i.e. what Peter Strawson (1962, 6) has called “reactive attitudes”. For Darwall (2006, 17) all such reactive attitudes implicitly involve a demand.

In addressing demands, or reactive attitudes to one another, we take ourselves to be members of the “moral community” (cf. P. Strawson 1962, 17) and thereby imply that the person to whom we address this demand, or, more broadly, expectation, is capable of addressing it to herself (Darwall 2006, 17; Darwall 2017, 268).¹⁸³ This capability is a presupposition of such practices and it must include both a basic understanding of the practice of demanding others and some sort of motivation to follow these demands in principle (cf. Darwall 2017, 269; 275f.). These basic facts concerning understanding the practice and being motivated to follow it must hold regarding the speaker as well if she is to sincerely demand another to do something (Cf. Darwall 2021, 564). Authority, demands and second-personal reasons therefore involve some sort of symmetry between speaker and addressee.¹⁸⁴

Considering morally relevant cases, this account can build on the Kantian idea of dignity (cf. Darwall 2006, 14). This might be understood as an account of the second formulation of the categorical imperative: “Treat humanity in each person always as an end in itself and never only as a means” (Kant 1786, B 66f.),¹⁸⁵ or, as Darwall (2006, 21) puts it: “[t]here is [...] a form of reciprocal respect that is built into all second-personal reason-giving”. Most basically, we

¹⁸² I omitted the emphasis.

¹⁸³ See also Darwall 2006, 21f. on “Fichte’s point” and *ibid.*, 23 on Pufendorf’s point. Fichte’s idea of a “summons” may be usefully compared with Williams’s idea of blame, and perhaps other “reactive attitudes” as proleptic (see above, III.5.d.; Darwall refers to this idea in 2006, 60). I cannot develop this any further here.

¹⁸⁴ In Darwall 2021, endnote 6 Darwall qualifies this claim, saying that second-personal competence is only sufficient, not necessary for second-personal authority in order to account for the capacity of animals or human beings who lack second-personal competence to make valid claims on us. He seems not to take this as the standard case though which is why I will disregard it in this cursory treatment.

¹⁸⁵ Slightly altered in translation. The original is: “Handle so, daß du die Menschheit, sowohl in deiner Person, als in der Person eines jeden andern, jederzeit zugleich als Zweck, niemals bloß als Mittel brauchst.”

can understand this to mean that any person has the authority to address certain demands, e.g. not to be abused, to others and to thereby give them second-personal reasons for acting. This authority, so the Kantian story goes, consists just in the ability to make demands, or, what comes to the same thing here, to make laws, generally, that is to oneself and others (cf. Korsgaard 1996a, 143; Darwall 2006, 257ff.).¹⁸⁶

This might be a good account of robust reason-giving. The statement made by the speaker is one that has the structure of making a law, i.e. it installs conditions that determine whether some action accords to it or not. The motivation of the addressee that makes this into a reason or not is her ability to make reasonable laws herself. If she recognises the statement by the other as a reasonable law she could make to herself, it will constitute a reason for her to act accordingly. And only if this is the case, the account may go, does it constitute a form of robust reason-giving. It surely has no problems with Roman penal law or the counterexamples to Enoch's view, and even the intuition that, in following the other's demands, we let ourselves be guided by the other's will might be accommodated by this view. After all, you do act on the law the other made in fulfilling their demand (cf. Korsgaard 1996a, 140-145).

This Kantian approach faces the criticisms that have been generally leveled against the project of grounding ethical relations on reason, of which Williams's (1985, 60ff.) is an important example.¹⁸⁷ Instead of relying on such criticisms though, I will focus more closely on the matter at hand. Can this story really provide a satisfactory account of robust reason-giving?

Lévinas-inspired critics have maintained that it fails to account for the asymmetry that is implicit in situations involving a demand (cf. Crowell 2020). We have stated above that, for Darwall, any demand presupposes the ability of the addressee to address the demand to herself. For Kantian theories such an internal address is the model of normativity generally. "The thinking self has the power to command the acting self, and it is only its commands that can make the action obligatory." (Korsgaard 1996a, 165)¹⁸⁸ The, at first sight surely, asymmetrical relation of demander and demandee is relegated to the internal sphere of the addressee so that the

¹⁸⁶ Korsgaard deduces from this that we must acknowledge others to be reason-giving in themselves using Wittgenstein's private language argument (cf. Korsgaard 1996a, 136ff.). As Williams (1996a, 216) has noted, this comes quite close to what Lévinas might have in mind. Lévinas himself has been understood to develop Kantian ethical ideas further too (see Gelhard 2015, 233ff.). Korsgaard's treatment differs still from Lévinas's in that she takes there to be reason in each of us and it is only this that explains to us why we should take others in themselves to be reason-giving. Lévinas takes the experience of others being reason-giving to be prior (see the next page).

¹⁸⁷ There is one criticism it might reasonably hope to evade though. That is the Hegelian point that ethics should be grounded on what is implicit in concrete ethical relations rather than on pure reason (cf. Hegel 1821, § 135). However, see Peterson/Samuel 2021 for criticism of Kantian constructivism inspired by Hegel.

¹⁸⁸ See also Korsgaard 1996a, 104f; Tarasenko-Struc 2020, 82.

relation between the actual human beings demanding and being demanded can stay symmetrical. But this would mean that demands by others were not possible if we did not already command ourselves. Contrary to what speech seems to imply, the paradigmatic model for telling someone what to do is then not a relation between two people but a relation of a person to herself. And isn't that just incredible?¹⁸⁹ As Lévinas (1961) seeks to show, the founding experience of such a command, i.e. the experience by reference to which we understand what a command is, necessarily involves a relation to someone else and the asymmetry of it lies precisely in the difference of one person to the other.¹⁹⁰ As Crowell (2020, 11) puts it: “I *learn* what the normative *is* through the other's command.”¹⁹¹

This does not mean that social relations characterised by reciprocity and symmetry are impossible. According to Lévinas (1961, 222f.; 241f.), experiencing the other's command initiates us into acting for reasons and giving reasons. This is so because the command requires that we justify what we do in the face of the other (cf. Crowell 2020, 13).¹⁹² Accordingly, we respond to the command coming from outside of us by figuring out what would satisfy it and thereby develop the capacity to consider and act for reasons. This seems quite strong. Surely, a human child set out in the wild could act for reasons and how would Lévinas account for that? Luckily, we do not need to engage with this any further. As the argument above shows, the entire discussion rests on the presupposition that normativity generally is of the type of conforming to a command, or, what might come to the same thing, a matter of obligations. As Williams (1985, 69-71; 213f.) has forcefully argued, this is a very poor picture of normativity and of ethics. Rational creatures are neither always “in the business of making rules” (Williams 1985, 69) nor are they always trying to obey some command.

¹⁸⁹ Tarasenko-Struc 2020, 83f.; 87 makes a similar point about Korsgaard's account of obligation and authority. Wouldn't it be outrageous to assume that we would owe rescuing someone to ourselves primarily rather than to the other? Relatedly, Williams (1996b, 216f.) responds to Korsgaard that it is not clear how a duty of respect for others could be deducted from the Kantian picture of the reflective self as the source of normativity.

¹⁹⁰ A difference that he explains with reference to Descartes's concept of God (cf. Lévinas 1961, 39-45). Arguably, Lévinas speaks mainly about the “other” and not demands, or requests. The “other” however is conceptualised as essentially giving orders (“commandements”, Lévinas 1961, 220). I cannot explore in how far Lévinas is really speaking about the same things about which we are speaking here and it might be that I partly misconstrue his work simply by taking it up here. I think it is nevertheless justified to include his ideas in this work because the main concepts Lévinas uses are arguably very relevant here and may enrich our thought. I am grateful to the participants of a master seminar held by Hans Bernhard Schmid at the University of Vienna in summer 2021 for useful discussion on this point.

¹⁹¹ Emphasis in the original.

¹⁹² The condition in which we are face the other is also characterized by Lévinas as “shame” (cf. Lévinas 1961, 82f.; see Crowell 2020, 15f.). It would be interesting to compare this account of shame with Williams's (1993, endnote 1). Unfortunately, I can't do this here.

Instead, what I think we should take away from this discussion is that it shows how the Kantian reconstruction of robust reason-giving misses a prime element about it. Even if certain of such cases could be analysed on the model of (ab-)using some internal capacity to demand things of ourselves, it couldn't be that all cases could be analysed in these terms. That would be putting the cart before the horse. Even though acting for reasons may not be, addressing demands is something that we do primarily to one another and not by ourselves. And, indeed, the child set out in the wild might act for reasons but it is much less certain that it will be addressing demands and making rules for itself in the same way we, indeed, often do.¹⁹³ If this is the account we should give of some cases of robust reason-giving, it could not be the account for all of those cases. So, there are at least some cases of robust reason-giving for which we still need a story. And perhaps, if we have such a story, we might not need the other one any more.

Therefore, another internalist account has failed to give us a satisfactory account of robust reason-giving. What may we conclude from this? The discussion has certainly not shown that it is conceptually impossible to give an internalist interpretation of robust reason-giving. I doubt that showing this would be possible. But, it should have made us attentive to the fact that they present a class of social situations in which reasons play a large role, that require some account in order to understand them, and for which internalism has no straightforward account to offer. This is sufficient reason to go on and look elsewhere for such an account. Indeed, perhaps, a concept of external reasons may be useful in providing one.

4. Intersubjectivist Objections to Internalism

An externalist account for robust reason-giving is under pressure on two fronts. It needs not only to show how it can explain the normative force of acts of robust reason-giving. It also needs to demonstrate that its concept of external reasons can satisfy the conditions set out in part II, i.e. it must meet Williams's challenge. In this chapter, I will examine whether intersubjectivist objections to internalism can deal better with this challenge than their objectivist counterparts. I will argue that two intersubjectivist proposals fail to do so but that their failure is instructive for what may be a consistent concept of external reasons able to meet Williams's challenge.

¹⁹³ Intending something as a way of giving oneself reason to do it may be a prime example of this (see e.g. Verbeek 2007). It is not implausible that the child set out in the wild would not intend things in that way.

a. Desire-independent reasons

I will begin with John Searle's concept of "desire-independent reasons" (cf. Searle 2001, 169ff.). Suppose, you go into a bar and order a beer (for the example cf. Searle 2001, 186f.). The time then comes for you to leave. Searle asks whether it depends on your subjective motivational set whether you have a reason to call the waiter and pay the beer. It seems rather obvious that it does not. Rather, in ordering the beer you have committed yourself to paying it afterwards. That is just what it *is* to order a beer (see. Searle 2001, 184). It would be rational to recognise this and, consequently, to be motivated to pay the beer (see Searle 2001, 189f.). What is recognised here would therefore be an external reason.

Recall that Williams's argument for internalism proceeded by issuing a challenge. Any account of reasons must show how (P1') and (P2) come out true under it. These two read:

(P1'): Any reason must be able to explain the action for which it is a reason if the agent is reasonable.

(P2): All actions are explained by reference to the agent's actual motivations.

Moreover, it must prove to be a satisfactory account of reasons in other ways. Specifically, it must first show us how this reason could explain an action *by the agent for whom it is a reason* and not by someone else, such as an ideal agent type (see above III.2. regarding McDowell). Second, it must give us an explanation of why some fact is a reason at all (see above III.3.c.: this is the requirement imposed in the first of the three arguments against Parfit). That is, it must provide us with a schema for a satisfactory answer to the normative question: Why should I do that?

For explaining (P1'), Searle (2001, 177; 190f.) draws on the idea of responding to a reason which we considered in discussing Parfit's objection to internalism. That one has ordered a beer and that this act has the normative significance of speaking in favour of paying for it afterwards is a fact. If one recognises this fact or, more precisely, its normative significance, one will be motivated to act accordingly. On this picture, being motivated to act accordingly is what it is to recognise the normative significance of a fact. Parfit would call this (correctly) responding to the reason to pay the beer given by the fact that one has ordered the beer (see above III.3.b.).

This account has no difficulty with (P1'). The reason to pay, i.e. having ordered the beer, may explain the paying if the agent is reasonable because recognising the normative significance of that fact just is what it is to be reasonable on this picture (see above II.1.). Moreover, it may explain this action in the sense of showing how it is justified from the perspective of the agent

(see above II.2.) because, if the agent comes to recognise that reason, she will acquire the motivation to act accordingly without changing in such a way as to make her into a different person. This motivation is a motivation to follow a reason the agent has and it is what explains why the action seems to be a good idea to her (see above II.3.b.).

If this account shares the strengths of Parfit's, it may share the weaknesses too, however. This may be examined by asking the normative question, once again. The agent may ask why her ordering a beer should be a reason to pay for it. Indeed, it is not so seldom that people ask themselves or others why they should do what they promised to do. Searle tells us that the ground for this reason is the voluntary binding of one's will in carrying out an institutionally defined speech act such as ordering something in a restaurant or promising to do something (cf. Searle 2001, 198f.). This binding is possible because it is a manifestation of one's freedom (Searle 2001, 202; 209f.).

But, this reason for the ordering of the beer's being a reason is hardly more convincing as a reason than the fact that one has ordered a beer itself (cf. above III.5.c.). Indeed, it is possible to argue that something like consistency of the will over time, which seems to be implied by the idea of voluntarily binding one's will for some later time, is not such a desirable thing after all.¹⁹⁴ I see three possibilities for Searle to answer the doubts about the normativity of freely binding one's will.¹⁹⁵

First, Searle may argue that binding one's will in promising and similar acts just is something normative and that there is no further explanation for its normativity.¹⁹⁶ But, if this is so, he will leave skeptics regarding the normative nature of binding one's will unsatisfied. More generally, he will leave everyone unsatisfied as to the quest for an explanation for that kind of normativity. Just as with Parfit, we will just be left with the observation that such things are normative for us and that's it. This implies that the question "Why should I do what I bound myself to do?" does not really have any sense and this contravenes our sense of this being an important question. In short, the problems we found with Parfit's concept of a reason (see above III.3.c.) apply here as well.

¹⁹⁴ For such an argument see G. Strawson 2012 turning against narrative theories which bear some resemblance to Searle's account in this regard.

¹⁹⁵ The same question could be raised regarding other accounts of commitment, cf. e.g. Gilbert 2018, 161ff.

¹⁹⁶ This may be indicated by his emphasis that promises involve desire-independent reasons "by definition" (cf. Searle 2001, 193; see also *ibid.*, 170; 173; 209).

Second, Searle may argue that the ability to bind one's will and to recognise this binding as a reason for action in the future is part of a principle of temporally extended rational agency.¹⁹⁷ Temporally extended rationality may require some sort of consistency with one's projects and commitments and the recognition of a promise as a reason for action is the recognition of this requirement regarding the case of promises. This surely is a possibility for Searle. His position would then be internalist however. Insofar as it is included in what it is for a human agent to be rational, it will be included in the notion of sound deliberation just as trying to be well-informed is included. Searle's position would then be akin to Kantian approaches to the extent that it proceeds on a thicker notion of what it is to be rational to provide us with a more substantive notion of the agent's reasons while still remaining internalist.¹⁹⁸

Third, Searle may argue that binding one's will in promising and similar acts derives its normativity from the institutional forms it uses.¹⁹⁹ It is in this third interpretation that Searle's objection will be intersubjectivist, properly speaking. It should be clarified right away that what is meant here is not that the institutional forms of promising are normative in that they provide a reason for everyone to use them but that they are normative in that they provide everyone who uses them, i.e. who, as a matter of fact, gives a promise, with reasons for a certain course of action. The problem with this third approach is that it would need to explain why anyone should have reason to respect the institutional forms given by a society.²⁰⁰ As it stands, it is therefore incomplete.

One such explanation may proceed on the idea that such respect is expected by other people in that society.²⁰¹ The explanation of why something is a reason for the agent to do something would then centrally invoke attitudes, e.g. expectations, of others. This approach has some initial plausibility if one has accepted internalism so far because here too reasons would depend on attitudes. It is only the subject of these attitudes that would change. Of course, this idea

¹⁹⁷ Such a reading may be suggested by Searle's (2001, 202; 209f.) claim that what makes commitments into reasons for action is that they are a "manifestation of the freedom" of the speaker (see also Searle 2001, 176).

¹⁹⁸ Regarding the Kantian version of internalism see above II.5.a. Searle (2001, 190-193) criticises Kant for requiring an extra desire in order to motivate someone to fulfill their duty.

¹⁹⁹ This reading goes against the text as Searle (2001, 198; 204) describes the institution of promising as the "tool" for creating desire-independent reasons. The reason for keeping the promise is supposed to derive from freely binding one's will using this institution. I still set it out here because it is an interesting way in which Searle could have developed his idea, and, for Searle, the institution of promising seems to be a structure that is necessary for anyone to promise anything (Searle 2001, 204). This may suggest that it is not just a random tool but that it may be a necessary feature of explaining desire-independent reasons for action.

²⁰⁰ Compare the Chess example above II.3.b.

²⁰¹ Again, Searle (2001, 178) only finds this idea plausible only against the backdrop of the agent freely having bound one's will.

would need to be spelled out more precisely in order to show how it might meet Williams's challenge for a plausible account for reasons.

Searle himself does not provide too much help in this regard. He does give an answer to the question of how institutional forms depend on attitudes of the people living in a society.²⁰² His explicit account of how these institutional forms can create reasons always resorts to the idea of freely binding one's will though (Searle 2001, 208-210). An intersubjectivist objection to internalism would have to derive the normativity of ordering one's beer more directly from the attitudes of others. If such an account may be given, Williams's challenge may be met without implicitly resorting to internal reasons.

We have seen that Searle's idea that ordering a beer is a reason to pay for it can be spelled out in three ways. The first two ways are unsatisfactory because they either fall prey to the same argument we have introduced against Parfit's objectivism, or they are, on closer looks, really just another version of internalism. An intersubjectivist objection would have to use the third way to spell this idea out. I will therefore inquire whether sense can be made of an explanation of reasons that makes reference not to the motivations of the agent but to attitudes of persons other than the agent.

b. Empathy and commitment

Recently, there have been some theorists who have put forward criticisms of internalism roughly on the basis of this idea.²⁰³ Of those cited, only Marion Hourdequin explicitly tackles Williams's internalism. I will therefore concentrate on her account in the following. I will proceed by setting it out and then providing criticism that should help us develop a better account of what external reasons on the intersubjectivist understanding might be.

The basic idea of Hourdequin's paper is that not only the ends of ourselves may ground reasons for action but also the ends of others. The reasons that derive from the ends of others are external as they do not rely on any ends or desires of the agent herself (cf. Hourdequin 2012, 404).²⁰⁴ They are part of sound deliberation not from the agent's motivations but from the motivations of someone else.

²⁰² Cf. Searle 1995; Searle 2001, 206-208.

²⁰³ Hourdequin 2012; Peterson/Samuel 2021; Crowell 2020.

²⁰⁴ Cf. also Peterson/Samuel 2021, 88; 99.

This is the core of any intersubjectivist proposal. It proposes that reasons may not only be something that depends on the motivational set-up of the agent herself but on the motivational set-up of other agents (see above II.5.c.). This proposal seems to have the strength that it is able to deal with the objections that we have brought forth against Parfit's objectivist account (see above III.3.c.). It allows for the possibility of an answer to the normative question in allowing an explanation for why even the most basic reasons are reasons for action – tying them back to the other's ends. Moreover, it retains the specific character of practical rationality that is implicit in Williams's notion of sound deliberation as it will use that notion itself in specifying how reasons depend on the motivations of others. Lastly, it shows how reasons depend on human existence and are not alien to it (see above III.3.c.). Indeed, it may charge internalism itself of having a sort of alienation as its consequence: “social alienation“ (Peterson/Samuel 2021, 84; 100f.), i.e. being disconnected from the peculiar normative situation of other people.²⁰⁵ However, there are also problems for such an account.

According to Hourdequin, we are confronted with the ends of others in empathy and collective intentionality. In order not to complicate matters unnecessarily, I will only speak about empathy in what follows. For Hourdequin, it is both a perception of the ends of other people and a source of motivation for us. The cognitive aspect consists in perceiving certain emotional, conative or other mental states of others, the motivational aspect consists in being able to *respond* to them by taking them as reasons for action (cf. Hourdequin 2012, 415f.).²⁰⁶

Thus, Hourdequin gives us both a story about what reasons are and about how they motivate us. This latter story may or may not satisfy (P1'). Let's see if it does.

There are two basic options for Hourdequin. She may either suppose that sound deliberation from all desires other people have may provide reasons for the agent. Besides the intuitive problems one may have with this option, it would then be hard to see how these reasons will entertain a specific relation to the agent (see above III.2.). After all, the desires other people have are the same for everyone and, consequently, everyone would have the reasons which derive from their desires in sound deliberation. This concept of reason could not make us understand how some action done for those reasons made sense specifically for the agent.

²⁰⁵ For a systematic account of “alienation” see Jaeggi 2005.

²⁰⁶ For a more detailed discussion of empathy see Zahavi 2014, part II.

Or, Hourdequin may, as she actually does, restrict the desires of others that may explain reasons for the agent in some way. For instance, Hourdequin (2012, 414) mentions the suffering of others as desires that are eligible for so explaining reasons of the agent.

There are, again, two ways to introduce such a restriction. The most obvious option would perhaps be to conceive of empathy as a motivational state that actually has some content, i.e. to relieve others from suffering. The reason why only some desires of others may be at the background of reasons for the agent is that they are of the kind which allow motivation through empathetic access, e.g. suffering. Then, the explication of what it is to be a reason does depend on a motivation of the agent herself: empathy. These reasons too will then depend on sound deliberation from the agent's actual motivations, taking a detour through desires of others. They will therefore be *internal* reasons.

This is, of course, a real option for giving an account of reasons someone may have for helping others. Right now, however, we want to see whether there is an *externalist* account which could meet Williams's challenge. In this context, this option must be discarded.

The second way to introduce such a restriction would be to allow morality to play some part in determining which desires of others are able to ground reasons for the agent. There would then be a substantive normative backstory which tells us which of the motivations of others provide reasons for action for us. But, to the extent that such substantive considerations determine which of, principally eligible, motivations of others, can actually ground reasons for us, we will be unable to ask why some fact should be a reason for us. All that we can do to justify such a substantive restrictions is appeal to principles of morality and for these it is unclear why they should be normative for us in the first place. That is, we cannot meaningfully ask the normative question about them. Consequently, in some cases, this question would have no meaning after all, which we held to be untenable (cf. above III.3.c.).

It follows that Hourdequin's account lacks the resources needed for giving a defensible externalist account. It fails to provide this because it fails to establish a stable third intersubjectivist option to internalism and its objectivist opponents. Hence, we need to reject her account and look somewhere else.²⁰⁷

²⁰⁷ Jennifer Hudin (2008) proposes what seems to me to be an intermediary position between Searle and Hourdequin. She introduces the notion of a moral sense that makes us perceive social commitments as obligatory without reflecting on this practically in Williams's sense. I think that this proposal would really still be internalist. The moral sense may count as a motivation in Williams's sense as it has the content of responding to social commitments specifically. That actual reflection from this sense is allegedly not possible does not mean that it could

c. Conditions for a good externalist account of robust reason-giving

Before doing so, it will be good to draw insights from the discussion of the two preceding intersubjectivist objections to internalism in the form of a plan for constructing an intersubjectivist account of external reasons. I will do so in setting out the central idea that an intersubjectivist account may pursue and two conditions for such a pursuit to be successful.

The central idea is that there is a concept of reasons that involves reference not to a motivational state of the agent but to attitudes of a person different from the agent.

The two conditions are as follows:

(C1): The concept of a reason must include and explain some restrictions on what counts as a reason given certain attitudes of others.

(C2): The account must be supplemented by pointing us to an attitude of the agent that (a) counts as responding to such reasons; (b) whose motivational aspect is not used to explain restrictions on what counts as a reason given certain attitudes of others and (c) does not involve the agent changing too much.

The central idea is the same as that proposed by Hourdequin. (C1) is a need also felt by Hourdequin. If it were not respected, sound deliberation from any desire, feeling of duty of whoever would provide reasons for the agent to act. Everyone's reasons would then be the same because everyone would derive reasons from all of their respective others' motivations and there would be no way in which we might say that reasons may make us see how an action makes sense specifically to the agent (see above IV.3.b.).

(C2a) incorporates the insight of how Parfit, and Searle, plan to explain how external reasons can satisfy the explanation condition, or, more specifically, (P1'). That is, this attitude must be defined by (a) being a correct perception of reasons that exist independently of it and by (b) being a motivating state of the agent that can be used to explain the agent's actions.

(C2b) clarifies that this attitude qua motivational attitude cannot be used to explain restrictions on what counts as a reason given the perceptual part. If it were, the account would be internalist (see above IV.3.b.).

(C2c) clarifies that this attitude must have a limited impact on the agent. More precisely, it cannot amount to something like a conversion in McDowell's sense. The specific identity of

not make an agent act upon sound deliberation, given that Hudin (2008, 246) allows for the individual disregarding her moral sense upon reflection from other motivations.

the agent must be preserved throughout the change from her not having the attitude to her having the attitude, such that the possibility that she have this attitude and be motivated by it can be recognised as hers (see above III.2.).

5. An Externalist Account of Robust Reason-Giving

I will now present my own account of what external reasons might look like. I will do so in first proposing a starting account of external reasons and then modifying it such that it meets the conditions set out above (IV.4.c.). In doing so, I suggest that this concept of an external reasons can meet Williams's challenge. Lastly, I will add some comments on how it could be used to solve the problem of robust reason-giving.

One caveat should be addressed first. The relationship between robust reason-giving and external reasons is much closer than the, correlative, relationship between internal reasons and epistemic reason-giving. I think that this should hardly be a surprise, given that the intuitive account of robust reason-giving was that it is a speech act in which reasons are not only cited but created. This close relationship should make us careful to avoid harmful circularity. We may be tempted to pass something that invokes robust reason-giving for an explanation of external reasons and then move on to use that notion of external reasons to give an account for robust reason-giving.

a. External reasons

Here, I will set out a first interpretation of the central idea of intersubjectivist objections to internalism in order to have a basis for dealing with the challenges of Williams's argument and the problem of robust reason-giving. This basis is:

- (E) A has external reason to X only if A could arrive at the conclusion to X by soundly deliberating from the actual motivations of someone else, B.

This statement includes two amplifications of the central intersubjectivist idea, corresponding to Williams's concept of internal reasons. First, the attitudes of the other that might ground reasons for the agent are motivations. Second, the conclusion to act on the basis of these motivations must be mediated by sound deliberation of the agent. These amplifications are necessary in order to explain the normative force of external reasons. The attitudes of someone other than the agent are no more something normative in themselves than the agent's attitudes.²⁰⁸ The

²⁰⁸ Cf. above III.3.a.

normativity must therefore come in somewhere else, and, given that we cannot accept an indefinable notion of reasons à la Parfit, the only available option seems to be sound deliberation. Sound practical deliberation in Williams's sense, in turn, could not proceed from any other attitudes than those that provide practical problems, i.e. motivations (see above II.3.a.).

Given (E), we can, paralleling the definition of an internal reason, give a preliminary definition of an external reason.

(ER) A fact is an external reason for A to X only if it could come up in A's sound deliberation based on someone else's, B's, actual motivations and potentially leading to A doing X.

This definition includes two more ideas that should be uncontroversial. First, it takes reasons to be facts (see above II.2.d.). Second, in speaking of "*potentially* leading A to doing X", it alludes to the possibility that there may nevertheless be stronger reasons for A not to do X. It should be noted right away that it is consistent with those stronger reasons being internal.

b. Meeting Williams's challenge

This account of reasons must meet (C1) and (C2) if it is to be acceptable. As it stands, it does not do so. It violates (C1) in taking every motivational attitude of everyone to ground reasons for action for any agent. Any agent will have (almost) the same external reasons then and there is hardly any point in saying that they could make actions of the agent intelligible to us in the way that (P1') as it was discussed in relation to McDowell (above III.2.) requires. It violates (C2) in not saying anything about the attitude that is supposed to help us explain A's acting for external reasons by providing an adequate motivation of her to do so. Consequently, I propose two ramifications of (E) and (ER).

These ramifications require, first, that the motivation applies to the agent and, second, that it is expressed to the agent. Both of these ideas will be taken in a minimal sense that I will explain shortly.²⁰⁹ There may be a larger sense of these ideas that already includes the concept of (external) reasons. Thus, for some potentially normative entity, e.g. a law, to apply to the agent, may be for it to give her a reason for action, as opposed to others to whom no such reason is given. For some potentially normative entity to be expressed to the agent might just mean that it is given to her in the way robust reason-giving works (see below IV.6.; though, usually, we may speak of "addressing" the agent here).

²⁰⁹ I am not trying to capture the common understanding, or something like the essence, of these ideas. I want to spell out ideas that are closely related to what is usually called applying and expressing and that are helpful here.

First, let's consider the idea of what it is for a motivation to apply to the agent. Here, I take it to mean only that the agent is part of a full specification of the content of the motivational attitude of the other. As an example, take Anna's desire to have dinner with her girlfriend. This desire "applies" to her girlfriend in the sense that she is part of its content and therefore fulfilling the desire would require her to do something.

The reason for amending (E) and (ER) using the idea of application specifically is that, just as the motivations of the agent herself, such motivations already include everything that is needed to derive actions of the agent from them. This is because the practical questions such motivations present already refer to the agent. The agent's own motivations may be considered under the form: "How can I act such that ...?" Motivations of other agents that apply to the agent may be considered under the form: "How can I and A (i.e. the agent) act such that...?" A simple desire of someone else, such as the desire for ice cream, contains no such reference and something else would be needed to derive actions of the agent from them, for instance the agent's motivational set containing a desire to help.

If the motivations that may ground external reasons for action are limited to motivations that apply to the agent in this sense, (C1) will be met. These motivations are a subset of the motivations mentioned in (E) and they are a subset that will be different for every agent.²¹⁰ This should meet the requirement that reasons derived from such motivations may explain the agent's action in telling us what the point of the action would be for the agent specifically. In principle, this explanation does not need to hold for any other agent and hence it will concern the agent specifically.

The idea of expression is a little more complicated. Here, I take it to mean that the motivation in question must be brought to the attention of the agent in an action that is explained partly by an internal reason that derives from this very motivation. It will take some explanation to see how this helps in meeting the requirements set out in (C2).

Remember that (C2) required that the externalist point us to an attitude which may show how external reasons can explain actions and satisfies conditions (C2a-C2c). Accordingly, this attitude must be a correct perception which is capable of motivating the agent, its motivational part must not impose any further restrictions on what may count as a reason and it must not involve

²¹⁰ Of course, it might be the same for some agents but this should be no difficulty. The point is that it is not principally the same and it will therefore be an interesting discovery if it were actually the same for two given people.

a complete change of the agent (see above IV.4.c.). If we can find an attitude which satisfies these conditions, (C2) will be met.

The idea of expression is important because the attitude I propose for satisfying (C2) is only suited to correctly perceiving attitudes which are expressed. This attitude is the attitude of what I will call “understanding”.²¹¹

Understanding may be defined as the “internal” perception of mental states of someone else. Remember that we defined motivation above as a first-personal concept. It is the consciousness of a practical problem that may be framed as a question about what to do (see above II.3.a.). There are two ways in which one can relate to such an attitude. One may know that it exists and that it has certain properties, consequences etc. The other is to “live through” it. It consists in taking it seriously in the sense that the problems posed by it are recognised as problems and facts that speak in favour of one solution rather than the other are recognised as reasons for this solution.²¹² Following H.L.A Hart’s (1961, 89-91) discussion of two point of views regarding law, I will call the first way to relate to such attitudes the external point of view and the second way the internal point of view on them. Confronted with the motivational attitudes of someone else, we may not only take the external point of view on them. We may also perceive them in the sense that the perception includes living through these attitudes by ourselves. This is a perception that takes the internal point of view of what is perceived: an “internal perception”. This is what I propose to call “understanding”.²¹³

This attitude satisfies the conditions set out in (C2). It meets (C2a) because it is a cognitive attitude that gets its standards of correctness from its object: the motivation of the other. This means that the understanding of an attitude of someone else can be right or wrong depending on whether it conforms to this attitude. It is also motivating insofar as it is a correct perception of its object. It is not just that, in understanding, the one who understands knows what properties

²¹¹ Descartes introduces a perhaps similar concept of “grasping” something as opposed to simply knowing it, see Descartes 1630, 25 (AT I, 152).

²¹² See also Williams’s discussion of two different ways to use thick concepts (Williams 1985, 157). It has been remarked that there is tension between the internal reason thesis and the idea of thick ethical concepts (Heuer 2012; Wiland 2013). This cannot be taken into account here because I don’t have room to introduce the idea of thick ethical concepts. I do think that there are interesting connections to be drawn between these discussions and my own.

²¹³ It should be noted that the very possibility of such a perceptual attitude with regards to the other’s motivations depends on the result of debates in social cognition and on different interpretations of how we understand others quite generally. Straightforwardly, it may resemble “simulationist” views but I am confident that it would also be (in fact, more) consistent with different conceptions of “direct social perception”. For a summary of the different positions in the field and a defense of a view of direct social perception see Zahavi 2014, part II and Gallagher 2020, ch. 4-7, espec. ch. 6. I cannot follow this through any further here. I must therefore take the defensibility of the view regarding social cognition implied by my account for granted.

or consequences the attitude in question has. Understanding is taking the internal point of view on that attitude and hence it *is* living it through. Living through a motivational attitude means taking it as a practical problem and facts that speak in favour of one way of resolving it as a reason. And this just is what we have called having a motivation and hence being motivated (see above II.3.a.). Consequently, the perception which constitutes understanding is also motivating – given that an action by the agent is derivable from this motivation. This is the case if the other’s motivation applies to the agent in the sense discussed above.

Understanding also satisfies (C2b) because its motivational aspect does not imply any restrictions on what counts as a reason. Any attitude of understanding a motivational act is *eo ipso* motivating. Lastly, it satisfies (C2c) because the agent need not become someone else. She especially need not come to be identical to the other agent, or take over all her other motivations. All that is required is that she views the practical problem that constitutes the motivation of the other which is the object of the act of understanding first personally, i.e. that she takes the internal point of view on it.

Now, we must explain why this attitude, in connection with (P1’), requires that the motivation must be expressed to the agent. Remember that (P1’) requires that a putative reason must be able to explain the agent’s action. (P2) says that it could only do so through a suitably connected motivation of the agent. The motivation needed is inherent in the attitude of understanding the other’s motivation. Hence, if understanding a motivation is impossible, the reason based on this understanding will not be able to explain an action of the agent and it will not satisfy an important requirement on what it is to be a reason.

It seems to be a suitable requirement for anyone to understand the mental states of someone else that they are in some way expressed. That is, absent telepathy, there must be some outward, bodily indications of these mental states. Most clearly, such expression is given when there is an action of the other that could only be explained using an internal reason of the other which is based on the expressed motivation. This action constitutes an outward perceptible thing which derives its meaning from the motivations which have led to it.²¹⁴ The notion of expression is meant to introduce the requirement that only motivations that are in such a way implied by a perceptible thing could ground reasons that are able to explain the agent’s action. The changes that would be needed for reasons that derive from motivations of others which are not yet expressed are too substantive for them to count as the actualisation of a reason’s ability to explain

²¹⁴ On the close relation between an action and the reasons that have led to it see above III.3.a.

action (see above III.1.). Therefore only motivations of others that are expressed may ground reasons for action for the agent.

The final account of external reasons I wish to propose therefore adds the ideas of application and expression to the earlier accounts.

(E'): A has external reason to X only if A could arrive at the conclusion to X by soundly deliberating from the actual motivations of someone else, B, which are expressed and apply to A.

(ER'): A fact is an external reason for A to X only if it could come up in A's sound deliberation based on someone else's, B's, actual motivations, which apply and are expressed to A, and potentially leading to A doing X.

The foregoing should be enough to show how this account of external reasons meets Williams's challenge. It may explain A's doing X, because it is part of the explanation of why A may be motivated to doing X. The actualisation of its ability to explain A's doing X may be characterized as being reasonable because it consist of a correct perception and sound deliberation thereupon. Such a reason may explain A's doing X specifically because this reason is a reason for her and it need not be a reason for anyone else: Although it may not be an expression of her values and desires, it is an expression of her practical situation and not just of what an ideal agent would do in her circumstances. Lastly, the normative question can always be asked about external reasons because there may in principle always be an answer that may say how a given fact satisfies (ER') or does not do so.²¹⁵

Perhaps, an example may make this a little clearer. Suppose that you come by a cyclist who has fallen, who is obviously injured and who expresses her pain in shouting for help. Besides whatever internal reasons you may have for rushing to help, calling an ambulance or similar things that are helpful in this situation, you also have an external reason to do so: roughly, that these actions will help the person with her injury. This fact is an external reason for you to help because it would favour helping in sound deliberation starting from the other's motivations – her desire to stop the pain – this desire is expressed to you through her shouting for help and it applies to anyone who is in a position to help (the cyclist wants people who can help to help), which you are ex hypothesi.

²¹⁵ This implies that the other's motivation is not itself a reason for acting for the agent. It is only part of the explanation of an external reason of hers. This resonates with our description of the intuitive normativity of demands, requests, etc. (see above IV.3.a.).

The explanation of an action such as calling an ambulance differs according to whether you do it for an external reason or for an internal reason. If you understand her situation correctly, i.e. if you comprehend the other's motivation to stop being in pain by first-personally living through it, you will eo ipso have that motivation and sound deliberation from that on may lead you to call an ambulance.²¹⁶ This is how the external reason may explain your calling the ambulance. The internal reason will involve no route through understanding the other's motivation in that sense. The motivation at issue will here probably be a desire to help others when they are in pain. All you need is the perception that the other is in pain and a very simple piece of sound practical deliberation to derive the conclusion to help her by calling the ambulance.²¹⁷

Another illustration for this concept of external reasons may be showing how it can be put to use in explicating an idea we use quite often: that sometimes our actions are someone else's in a particular sense (cf. Gilbert 2018, 71ff.).²¹⁸ For instance, we may read in history books that some palace was built by a specific emperor or king. It is never the case that these persons were the ones who actually did the physical labour of building the palace. Most often, they didn't even plan it. All they did was order others to do it. Apparently, that suffices for saying that the action of building is theirs in some sense, as opposed to the workers who actually got the job done.

Later in her discussion, Gilbert mentions an interpretation of this idea that she calls the "potential determination sense". Accordingly, "the co-creators of the joint commitment have *directly determined how a given party will act*, all equal, *if he acts in his capacity as a rational agent, properly setting aside his inclinations and the like*" (Gilbert 2018, 173).²¹⁹ The terms "joint commitment" and the idea of setting aside the inclinations of the agent belong to her theory of

²¹⁶ This does not mean that you feel the other's pain as if it were yours. As indicated above (III.3.d.), it is plausible that suffering from something is not equivalent to having an aversion to something. The pain and the motivation are two different things. And, really, it seems hard to believe, though not impossible, that you actually feel the cyclist's pain when you see her lying there. Presumably, your leg does not hurt (which is not to say that you may not be in other pain because seeing others be in pain makes you feel bad; but this is surely not the pain the cyclist is feeling). This is why I am saying that it is the motivation, not the suffering, which you (almost) share with the cyclist (you share it almost because for you it is a perception and for the cyclist it is not, so there still is a difference).

²¹⁷ The internalist may say that the first case is not special because you "understand" the other's motivation. It is special because you imagine to be in the other's pain and imagine what you would want in that kind of situation. This is then what you do. I am inclined to say that imagining to be in the other's place is an epistemic tool, and it is one that you could use both to understand the other's motivations and to know more about it. What it is you do with it, depends on whether you take the internal perspective on the other's motivations in the sense of seeing as a problem what she seems as a problem, or not. So, it could be part both of the description of a situation in which the agent acts for internal reasons and in which she acts for external reasons. I am grateful to Martin Niederl for raising this issue.

²¹⁸ See also above IV.3.a.

²¹⁹ Emphasis in the original.

how someone may gain the standing to demand an action from someone and need not concern us here. What is important is that one conception of the idea of an action belonging to someone other than the agent is that that other person has determined directly how that agent will act if she acts in her capacity as a rational agent. If the agent has actually acted that way, it will be the other person's in the "actual determination sense" (Gilbert 2018, 174). It is this sense that we alluded to in the above example of the emperors and their palaces.

The notion of external reasons just introduced is precisely what spells out these combined ideas. They are the kinds of reasons an agent may have that we can intelligibly hold to have been determined by another person. In determining the agent to have these reasons, that other person has, by tautology, also determined, to some extent, how the agent might act in her capacity as a rational agent. We can therefore say that an action the agent performed is someone else's if and only if it may be, or has been, performed for a reason that is external and based on the motivations of that other person. This idea is also what might be meant by "letting the other person's will guide our actions" or "acting for someone else's reasons and not our own" (see above IV.3.a. and Hanser 2015). We can also understand them as referring to actions done for an external reason in the sense captured by (E') and (ER').

c. Meeting the robust reason-giving challenge

We must still examine whether this notion of external reasons can give us an acceptable account of acts of robust reason-giving. I think that it can. Consider the following:

(RRE): A speech act is an act of robust reason-giving only if the speaker, S, intentionally expresses a motivation to the addressee, A, that applies to her and thereby makes her have external reason to perform some action.²²⁰

An act of robust reason-giving would consequently *be* the actions that are needed to explain an external reason, thereby validating the intuition that in such acts reasons are "created". Though, as we have already suspected (IV.3.a.), it *is* not the reason for action much like the motivations of the agent herself are not her internal reasons. They are only part of the explanation of their normative force.

²²⁰ Hanser (2015, 112-115) develops a similar account. Specifically, fulfilling another's request or demand is also conceptualised as following an end expressed in the command or request. I believe that he is trying to capture the very same idea. I have not incorporated his insights more fully into this chapter because I only learnt about his paper shortly before finishing this thesis. I hope to be able to compare his approach to the one set out here on a future occasion.

Robust reason-giving is not simply the act of expressing such a motivation though. It is intentionally so and the expression is intentionally made to specified addressees. This should capture the idea that robust reason-giving is a particular kind of action that people, at least in the normal case, do consciously. It also serves to exclude counterexamples such as audibly moaning in pain thereby expressing one's desire that someone come and help but not intentionally to anybody in particular. This moaning would be part of what constitutes an external reason to help but it would not count as an act of robust reason-giving, e.g. a request to someone else that they help.

I cannot go through all possible objections to such a view but it should be remarked that it has no problems dealing with the counterexamples to the other views we have considered in chapters 1-3 of this part. Making a statement that constitutes a change in the circumstances, such as saying "I love you" as a way of making a relationship more significant, is not directed at giving someone external reason for action but to modify the situation such as to make her have different internal reasons (see above IV.2.). Insulting somebody is so as well, unless the desire that the other person leave is expressed and it is intended to thereby give that person external reason for leaving. In that case perhaps, the insult may just be a disguised demand though (see above IV.3.a.). Lastly, this account has no problems with the intuition that we let ourselves be guided by the other person's will when we follow her demands or requests (IV.3.a.) – to the contrary, it allows to spell these ideas out further. Equally obviously, it has no problems with showing us how requests or demands are primarily addressed to us by others and not by ourselves (IV.3.b.).

It may be objected to this concept of robust reason-giving that it follows from it that anyone has the standing to demand or request anything from others as long as one has a motivation which applies to the others and which can be expressed to them. It would follow from this that I could demand anyone to fund my tennis career and thereby robustly give them a reason to do it. This is incredible (see above IV.2.).

The reply to this objection has two parts. First, (RRE) is a necessary, not a sufficient condition for a given speech act to be an act of robust reason-giving. A further condition may be that the speaker has the standing to demand the object of the motivation which is expressed from the addressee. What it is to have this standing, is, as I pointed out above (IV.1.), a different question.

Second, regarding requests the case is not as simple as the previous paragraph suggests. I submit that it is not so implausible that my requesting you to fund my tennis career is an external reason

for you to do so. You just have overwhelming internal reasons not to. But, contrary to commands, requests are not peremptory (cf. Gilbert 2018, 60).²²¹ This means that you are not rationally required to disregard your own internal reasons to any extent based on my request. And, you certainly have much stronger internal reasons to save up the money you would spend on my tennis career or to spend it on something else. If you don't, you might actually give me money for buying tennis equipment etc. and this would be best explained by reference to my requesting you to do so. This would certainly say a lot about your internal reasons (or, rather, about their absence) though.

6. Robust Reason-Giving and Rational Persuasion

I cannot go as far in depth about the part of the theory of rational persuasion which concerns robust reason-giving as regarding the part concerning epistemic reason-giving (see above III.5.). In this chapter, all I want to do, is emphasise that including it in a theory of rational persuasion does not mean that it has the same ethical standing as advice (see above III.5.c.). Indeed, pointing out how it is different from epistemic reason-giving, both in the act of reason-giving and in the reasons it gives, should make us attentive to the possibility that rational persuasion may be even more nuanced and normatively ambiguous than Williams suggested.

The reason why epistemic reason-giving is in principle a good thing to do does not apply to robust reason-giving. For epistemic reason-giving, the good-making trait was roughly that it furthered the addressee's practical freedom in that it helped her to effectively further some project or cause she endorses. This clearly does not apply to robust reason-giving. It is precisely not the addressee's causes or projects which are at issue but the speaker's. Indeed, the requests and demands others make on us often seem to be a hindrance to our practical freedom in that they keep us from considering what would be required for developing our own projects further. It may therefore be desirable not to act for the external reasons one may have. Indeed, it may be regarded as a sign of the strength of one's character to neglect demands or requests by others. This is commonly expressed by saying that someone knows "how to say no".

This does not mean that we should simply relegate robust reason-giving to the side of the bad ways of making others do things, i.e. what we have called "force" (see above II.1.a.). It may

²²¹ For James Lewis (2018, 6-8), they are "discretionary" which, roughly, implies a choice on the part of the agent of whether to fulfill them or not. See also Raz 1986, 36f. For a discussion of the "peremptory" quality of commands with reference to the works of Raz and H.L.A Hart, see Hanser 2015, 96-98.

achieve something that has been considered to be impossible on an internalist picture (cf. Peterson/Samuel 2021, 84; 100f.): it may save us from “social alienation”. Being attentive to our external reasons may allow us to escape being entrenched in our own practical stance and experience community and intersubjective exchange with others. It opens us to a world of different projects which give the facts of the world value and normative significance. And, this is surely something we consider to be ethically valuable.

We should conclude from this that the distinction between rational persuasion and force, which has seemed to be very clear and sharp at the beginning, is not clear and sharp after all. At least robust reason-giving is a form of rational persuasion which is neither good nor bad (or both good and bad). It is therefore not warranted to distinguish between the ethically valuable practices of reason-giving on the one hand and other, ethically ambiguous, ways of getting people to do things on the other. Although we have arrived at this conclusion through a criticism of Williams’s view, this is an idea he would surely have endorsed.

V. Conclusion

The results of the discussion of this thesis can be summarised in four statements: (1) there are two kinds of reasons: internal reasons and external reasons; (2) external reasons are not what the predominant objectivist critics of internalism have understood them to be – they are not just out there but they depend on motivational states of agents different from the agent who is, or will be, the subject of the action under consideration; (3) internal reasons and external reasons may help to define two fundamental types of rational persuasion: epistemic and robust reason-giving; (4) rational persuasion is not as ethically unambiguous as it may have seemed.

These results contain the two main theses which have defined our project in the introduction in (1)-(3). In part two, we have seen that there is a class of reasons for which a connection to the agent's motivations is a necessary condition and objectivist objections against this fail. This class of reasons defines an act of rational persuasion called "epistemic reason-giving". In part three, I have suggested both that there is reason for looking for a second type of reasons which are external in that, for them, a connection to the agent's motivation is not a necessary condition and that there is reason for thinking that such a second type of reasons is conceivable.

There are two obvious ways in which the thoughts presented in this text may be developed further. I have not set out a theory of rational persuasion here but only presented building blocks for such a theory. Both the account of epistemic reason-giving and of robust reason-giving may be further developed and systematised. Also, such a theory should contain a more detailed treatment of the ethical value of such acts. This may not mean providing a theoretical account of this value but presenting a more differentiated account of what we deem valuable about them precisely. Also, as mentioned multiple times already, the third part is presented as a proposal rather than as a definite argument. It would be necessary to deal with the arguments and concepts used and developed there much more intensely.

I am confident that, adequately incorporated in a more comprehensive theoretical account of reasons and rational persuasion, the results could be important for understanding both what it is to be an agent and important parts of our social life. Its ideas could be put to use in the explanation of behaviour, in (public) deliberation and in the analysis of (public) speech acts and communication. Of course, this is not so much a merit of this work as of the work which started the discussion on internal reasons: Bernard Williams's articles on the matter. But, perhaps I can be allowed to hope that the work done here can advance our understanding of these things a little more: if not in what it gets right, at least in its failures.

References

- Alvarez, Maria (2010): *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford: Oxford University Press.
- Arpaly, Nomy; Schroeder, Timothy (2014): *In Praise of Desire*. New York: Oxford University Press.
- Austin, John Langshaw (1962): *How to do things with words*. 2nd ed. Edited by J.O. Urmson and Marina Sbisa. Oxford: Clarendon 1975.
- Beaulieu, Gerald (2013): “Can Explanatory Reasons Be Good Reasons For Action?” In: *Metaphilosophy* 44 (4), 440-450.
- Berlin, Isaiah (1969): “Two Concepts of Liberty”. In: Isaiah Berlin (ed.): *Four Essays on Liberty*. New York: Oxford University Press, 112-172.
- Broome, John (2004): “Reasons”. In: R. Jay Wallace, Philip Pettit, Samuel Scheffler, Michael Smith (eds.): *Reason and Value: Themes From the Moral Philosophy of Joseph Raz*. NY: Oxford University Press.
- Brunero, John (2007): “McDowell on External Reasons”. In: *European Journal of Philosophy* 16 (1), 22-42.
- (2013): “Reasons as Explanations”. In: *Philosophical Studies* 165 (3), 805-824.
- Butler, Judith (1997): *Excitable Speech. A Politics of the Performative*. London/NY: Routledge 2021.
- Christman, John (2022): “Introduction. The Multiple Dimensions of Positive Freedom”. In: John Christman (ed.): *Positive Freedom. Past, Present, and Future*. Cambridge: Cambridge University Press.
- Crowell, Steven G. (2020): “Second-Person Reasons. Darwall, Levinas and the Phenomenology of Reason”. In: Michael Fagenblat and Melis Erdur (eds.): *Levinas and Analytic Philosophy. Second Person Normativity and the Moral Life*. New York: Taylor and Francis.
- Cunningham, Joseph J. (2021): “The matter of motivating reasons”. In: *Philosophical Studies*. Open Access. 1-27. <https://doi-org.uaccess.univie.ac.at/10.1007/s11098-021-01719-9>.
- Dancy, Jonathan (2000): *Practical Reality*. New York: Oxford University Press.
- Darwall, Stephen (1992): “Internalism and Agency”. In *Philosophical Perspectives* 6, 155-174.
- (2006): *The Second-Person Standpoint. Morality, Respect, and Accountability*. Cambridge/London: Harvard University Press.
- (2013): “Authority and Second-Personal Reasons for Acting”. In: Stephen Darwall: *Morality, Authority, and Law: Essays in Second-Personal Ethics I*. Oxford: Oxford University Press 2013, 135-151.
- (2017): “Morality, Blame and Internal Reasons”. In: Peter Singer (ed.): *Does Anything Really Matter? Essays on Parfit on Objectivity*. Oxford: Oxford University Press, 259-279.
- (2021): “Recognition, second-personal authority, and nonideal theory”. In: *European Journal of Philosophy* 29 (3), 562-574.
- Davidson, Donald (1963): “Action, Reasons, and Causes”. In: Donald Davidson: *Essays on Actions and Events*. 2nd edition. Oxford: Oxford University Press 2001, 3-19.
- (1969): “How is Weakness of the Will Possible?” In: Donald Davidson: *Essays on Actions and Events*. 2nd edition. Oxford: Oxford University Press 2001, 21-42.
- Deixler-Hübner, Astrid; Mitgutsch, Ingrid (2007): *Rechtlicher Schutz in Familie und Partnerschaft*. Mit einem Beitrag von Rotraut A. Perner. Wien: Facultas.
- Descartes, René (1630): “Letter to Mersenne, 27 May 1630”. In: *The Philosophical Writings of Descartes. Volume III. The Correspondence*. Transl. by John Cottingham, Robert Stoothoff, Dugald

- Murdoch, Anthony Kenny. Cambridge/New York/Melbourne: Cambridge University Press, 25-26.
- Enoch, David (2011): "Giving Practical Reasons". In: *Philosopher's Imprint* 11 (4), 1-22.
- (2014): "Authority and Reason-Giving". In: *Philosophy and Phenomenological Research* LXXXIX (2), 296-332.
- Fine, Kit (1994): "Essence and Modality: The Second *Philosophical Perspectives* Lecture". In: *Philosophical Perspectives* 8, 1-16.
- Finlay, Stephen (2006): "The reasons that matter". In: *Australasian Journal of Philosophy* 84 (1), 1-20.
- (2009): "The Obscurity of Internal Reasons". In: *Philosopher's Imprint* 9 (7), 1-22.
- Finlay, Stephen; Schroeder, Mark (2017): "Reasons for Action: Internal and External". In: *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/reasons-internal-external/#MotArg> (last access 06.06.2022).
- Forst, Rainer (2015): *Normativität und Macht. Zur Analyse sozialer Rechtfertigungsordnungen*. Berlin: Suhrkamp.
- Foucault, Michel (1984): "Die Ethik der Sorge um sich als Praxis der Freiheit". In: Michel Foucault: *Schriften in vier Bänden. Dits et Ecrits. Band 4*. Transl. by Hermann Kocyba. Frankfurt/M.: Suhrkamp 2005, 875-902.
- Gallagher, Shaun (2020): *Action and Interaction*. Oxford: Oxford University Press.
- Gelhard, Andreas (2015): „Skepsis und Schwäche. Versuch über Kant und Lévinas“. In: Alfred Bodenheimer and Miriam Fischer-Geboers (eds.): *Lesarten der Freiheit. Zur Deutung und Bedeutung von Emmanuel Lévinas Difficile Liberté*. Freiburg/München: Karl Alber 2015.
- Gilbert, Margaret (2018): *Rights and Demands. A Foundational Inquiry*. Oxford: Oxford University Press.
- Habermas, Jürgen (1981a): *Theorie des kommunikativen Handelns. Band 1. Handlungsrationalität und gesellschaftliche Rationalisierung*. Frankfurt/M.: Suhrkamp 1995.
- (1981b): *Theorie des kommunikativen Handelns. Band 2. Zur Kritik der funktionalistischen Vernunft*. Frankfurt/M.: Suhrkamp 1995.
 - (1992): *Faktizität und Geltung. Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*. Frankfurt/M.: Suhrkamp 1998.
- Hanser, Matthew (2015): „Doing another's bidding“. In: George Pavlakos and Veronica Rodriguez-Blanco (eds.): *Reasons and Intentions in Law and Practical Agency*. Cambridge: Cambridge University Press, 95-120.
- Hart, H.L.A. (1961): *The Concept of Law*. 3rd ed. With a Postscript. Ed. by P. Bulloch and J. Raz and with an introduction by L. Green. Oxford: Oxford University Press 2012.
- Hegel, G.W.F. (1821): *Grundlinien der Philosophie des Rechts oder Naturrecht und Staatswissenschaft im Grundrisse*. Ed. by Eva Moldenhauer and Karl Markus Michel. Frankfurt/M.: Suhrkamp 1986 [=Werke 7].
- Herstein, Ori (2017): "Understanding standing: permission to deflect reasons". In: *Philosophical Studies* 174 (12), 3109-3132.
- Heuer, Ulrike (2012): „Thick Concepts and Internal Reasons“. In: Ulrike Heuer and Gerald Lang (eds.): *Luck, Value, and Commitment: Themes from the Ethics of Bernard Williams*. Oxford: Oxford University Press.
- Hlobil, Ulf (2020): "Why you cannot make people better by telling them what is good". In: *European Journal of Philosophy* 28 (4), 986-996.
- Horkheimer, Max (1937): "Traditionelle und kritische Theorie". In: Max Horkheimer: *Die gesellschaftliche Funktion der Philosophie. Ausgewählte Essays*. Frankfurt/M.: Suhrkamp 1974, 145-200.

- Hourdequin, Marion (2012): „Empathy, Shared Intentionality, and Motivation by Moral Reasons“. In: *Ethical Theory and Moral Practice* 15 (3), 403-419.
- Hudin, Jennifer (2008): “The Logic of External Reasons and Collective Intentionality”. In: H.B. Schmid, Katinka Schulte-Ostermann, Nikos Psarros (eds.): *Concepts of Sharedness: Essays on Collective Intentionality*. Berlin/Boston: De Gruyter, 225-251.
- Jaeggi, Rahel (2005): *Entfremdung. Zur Aktualität eines sozialphilosophischen Problems*. Mit einem neuen Nachwort. Berlin: Suhrkamp 2016.
- Kant, Immanuel (1786): *Grundlegung zur Metaphysik der Sitten*. Ed. By Wilhelm Weischedel. Frankfurt/M.: Suhrkamp 1974 [Werkausgabe Bd. VII].
- (1788): *Kritik der praktischen Vernunft*. Ed. By Wilhelm Weischedel. Frankfurt/M.: Suhrkamp 1974 [Werkausgabe Bd. VII].
 - (1797): *Die Metaphysik der Sitten. Metaphysische Anfangsgründe der Tugendlehre*. Ed. by Wilhelm Weischedel. Frankfurt/M.: Suhrkamp 1977 [Werkausgabe Bd. VIII].
- Kelsen, Hans (1960): *Reine Rechtslehre*. 2nd ed. Wien: Verlag Österreich 2020.
- Kissine, Mikhail (2013): *From Utterances to Speech Acts*. New York: Cambridge University Press.
- Korsgaard, Christine (1986), “Skepticism about Practical Reason”. In: Christine Korsgaard: *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press 1996, 311-334.
- (1996a): *The Sources of Normativity*. Ed. by Onora O’Neill. Cambridge: Cambridge University Press.
 - (1996b): “Replies”, in: Christine Korsgaard: *The Sources of Normativity*. Ed. by Onora O’Neill. Cambridge: Cambridge University Press, 219-258.
 - (2008): “Acting for a Reason”, in: Christine Korsgaard: *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford: Oxford University Press, 207-230.
 - (2009): *Self-Constitution. Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Lear, Jonathan (2005): *Freud*. Abingdon: Routledge.
- Lefort, Claude (1979): “Formation et autorité: l’éducation humaniste”. In: Claude Lefort: *Écrire. À l’épreuve du politique*. Paris: Calmann Lévy 1992, 209-226.
- Lévinas, Emmanuel (1961): *Totalité et Infini. Essai sur l’extériorité*. Paris: Kluwer Academic 1990.
- Lewis, David (1973): “Causation”. In: David Lewis: *Philosophical Papers. Volume II*. New York/Oxford: Oxford University Press 1986, 159-213.
- Lewis, James H.P. (2018): “The Discretionary Normativity of Requests”. In: *Philosophers’ Imprint* 18 (20), 1-16.
- Lin, Eden (2020): “Future Desires, the Agony Argument, and Subjectivism about Reasons”. In: *Philosophical Review* 129 (1), 95-130.
- Lindeman, Kathryn (2020): “Conservative speech”. In: *Ratio* 33 (4), 243-254.
- Löschke, Jörg (2021): “Agent-Relative Reasons and Normative Force”. In: *Philosophia* 49 (1), 359-372.
- Lubin, Dean (2009): “External Reasons”. In: *Metaphilosophy* 40 (2), 273-291
- Manne, Kate (2014): “Internalism about reasons: sad but true?” In: *Philosophical Studies* 167 (1), Special Issue: Proceedings of the 2013 Bellingham summer philosophy conference, 89-117.
- Mantel, Susanne (2018): *Determined By Reasons. A Competence Account of Acting for a Normative Reason*. New York/Abingdon: Routledge.
- Markovits, Julia (2010): “Acting for the Right Reasons”. In: *The Philosophical Review* 119 (2), 201-242.
- (2014): *Moral Reason*. Oxford: Oxford University Press.

- (2017): "On what it is to matter". In: Simon Kirchin (ed.): *Reading Parfit. On What Matters*. Abingdon/New York: Routledge, 54-81.
- McDowell, John (1995): "Might there be External Reasons?" In: J.E.J. Altham and Ross Harrison (eds.): *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge: Cambridge University Press, 68-85.
- Milgram, Elijah (1996): "Williams' Argument Against External Reasons". In: *Noûs* 30 (2), 197-220.
- Mitova, Veli (2016): "Clearing space for extreme psychologism about reasons". In: *South African Journal of Philosophy* 35 (3), 293-301.
- Monti, Ezequiel H. (2021): "Against triggering accounts of robust reason-giving". In: *Philosophical Studies* 178 (11), 3731-3753.
- Moore, A. W. (2006): "Commentary on the text". In: Bernard Williams: *Ethics and the Limits of Philosophy*. Abingdon: Routledge 2006, 225-248.
- Morganti, Matteo; Tanyi, Attila (2017): "Can Reasons Be Propositions? Against Dancy's Attack on Propositionalism". In: *Theoria* 83 (3), 185-205.
- Nagel, Thomas (1986): *The View from Nowhere*. New York: Oxford University Press.
- Paakkunainen, Hille (2018): "Internalism and Externalism about Reasons". In: Daniel Star (ed.): *The Oxford Handbook of Reasons and Normativity*. Oxford: Oxford University Press, 143-169.
- Parfit, Derek (1997): "Reasons and Motivation". In: *Proceedings of the Aristotelian Society. Supplementary Volumes* 71, 99-146.
- (2011a): *On What Matters. Volume One*. Ed. and introduced by Samuel Scheffler. Oxford: Oxford University Press.
- (2011b): *On What Matters. Volume Two*. Ed. and introduced by Samuel Scheffler. Oxford: Oxford University Press.
- (2017): *On What Matters. Volume Three*. Oxford: Oxford University Press.
- Peterson, Christa; Samuel Jack (2021): "The Right and the Wren". In: David Shoemaker (ed.): *Oxford Studies in Agency and Responsibility. Volume 7*. Oxford: Oxford University Press, 82-103.
- Pettit, Philip; Smith, Michael (2006): "External Reasons". In: Cynthia Macdonald and Graham Macdonald (eds.): *McDowell and his Critics*. Oxford: Blackwell Publishing, 142-179.
- Rawls, John (1993): *Political Liberalism*. Expanded Edition. New York: Columbia University Press 2005.
- Raz, Joseph (1975): *Practical Reason and Norms*. 2nd ed. Oxford: Oxford University Press 1999.
- (1986): *The Morality of Freedom*. Oxford: Clarendon Press.
- (2015): "Normativity: The Place of Reasoning". In: Joseph Raz: *The Roots of Normativity*. Ed. and with an introduction by Ulrike Heuer. Oxford: Oxford University Press 2022, 72-93.
- Rosati, Connie S. (2016): "Moral Motivation". In: *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/moral-motivation/#HumVAnthum> (last access 06.06.2022).
- Sagar, Paul (2014): "Minding the Gap: Bernard Williams and David Hume on Living an Ethical Life". In: *The Journal of Moral Philosophy* 11 (5), 615-638.
- Samuel, Jack (2021): "An Individual Reality". In: *Inquiry*, ahead of print, 1-21. <https://doi-org.uac-cess.univie.ac.at/10.1080/0020174X.2021.1948445>.
- Scanlon, Thomas (1998): *What We Owe to Each Other*. Cambridge/London: Harvard University Press 2000.
- (2014): *Being Realistic about Reasons*. Oxford: Oxford University Press.
- Schraubroek, Katrien (2012): *The Normativity of What We Care About: A Love-Based Theory of Practical Reasons*. Leuven: Leuven University Press

- Schroeder, Mark (2007): *Slaves of the Passions*. Oxford: Oxford University Press.
- Schroeder, Timothy (2015): "Desire". In: *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/desire/> (last access 06.06.2022).
- Schueler, G.F. (1995): *Desire: Its Role in Practical Reason and the Explanation of Action*. Cambridge/London: The MIT Press.
- Schwarz, Wolfgang (2020): "Ability and Possibility". In: *Philosopher's Imprint* 20 (6), 1-21.
- Schweikard, David P.; Schmid Hans Bernhard (2020): "Collective Intentionality". In: *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/collective-intentionality/#TeaRea> (last access 06.06.2022).
- Searle, John (1969): *Speech Acts: an essay in the philosophy of language*. Cambridge: Cambridge University Press.
- (1995): *The construction of social reality*. New York: Free Press.
 - (2001): *Rationality in Action*. Cambridge/London: The MIT Press.
- Setiya, Kieran (2004): "Against Internalism". In: *Noûs* 38 (2), 266-298.
- (2014): "What is a reason to act?". In: *Philosophical Studies* 167, 221–235
- Shemmer, Yonatan (2019): "Subjectivism about Future Reasons or The Guise of Caring". In: *Philosophy and Phenomenological Research* CI (3), 630-648.
- Skorupski, John (2007): "Internal Reasons and the Scope of Blame". In: Alan Thomas (ed.): *Bernard Williams*. Cambridge: Cambridge University Press, 73-103.
- Sobel, David (2001): "Explanation, Internalism and Reasons for Action". In: David Sobel: *From Valuing to Value: Towards a Defense of Subjectivism*. Oxford: Oxford University Press 2016, 144-161.
- Strawson, Galen (2012): "We Live Beyond Any Tale That We Happen to Enact". In: *Harvard Review of Philosophy* 18 (1), 73-90, DOI: 10.5840/harvardreview20121814.
- Strawson, Peter F. (1962): "Freedom and Resentment". In: P.F. Strawson: *Freedom and Resentment and other essays*. London: Methuen 1974, 1-25.
- Tarasenko-Struc, Aleksy (2020): "Kantian constructivism and the authority of others". In: *European Journal of Philosophy* 28 (1), 77-92.
- Taylor, Charles (1989): *Sources of the Self. The Making of Modern Identity*. Cambridge: Harvard University Press.
- Teichmann, Roger (2019): "How should one live? Williams on practical deliberation and reasons for acting". In: Sophie Grace Chappell and Marcel van Ackeren (eds.): *Ethics Beyond the Limits. New Essays on Bernard Williams' Ethics and the Limits of Philosophy*. New York: Routledge, 209-221.
- Temkin, Larry S. (2017): "Has Parfit's Life Been Wasted? Some Reflections on Part Six of *On What Matters*". In: Peter Singer (ed.): *Does Anything Really Matter?: Essays on Parfit on Objectivity*. Oxford: Oxford University Press, 1-33.
- Thomson, Judith Jarvis (2001): *Goodness and Advice*. Ed. and introduced by Amy Gutmann. Princeton: Princeton University Press.
- Tubert, Ariela (2016): "Sound Advice and Internal Reasons". In: *Pacific Philosophical Quarterly* 97 (2), 181-199.
- Verbeek, Bruno (2007): "Rational Self-Commitment". In: Fabienne Peter and H.B. Schmid (eds.): *Rationality and Commitment*. Oxford: Oxford University Press, 150-174.
- Wiland, Eric (2012): *Reasons*. London: Bloomsbury Publishing.

- (2013): "Williams on Thick Ethical Concepts and Reasons for Action". In: Simon Kirchin (ed.): *Thick Concepts*. Oxford: Oxford University Press, 210-215.

Williams, Bernard (1972): *Morality. An Introduction to Ethics*. Canto Edition. Cambridge: Cambridge University Press 1993.

- (1973): "A critique of utilitarianism". In: J.J.C. Smart and Bernard Williams: *Utilitarianism for and against*. New York: Cambridge University Press, 77-150.
- (1979a): "Internal and External Reasons". In: Bernard Williams: *Moral Luck: philosophical papers, 1973-1980*. Cambridge: Cambridge University Press 1981, 101-113.
- (1979b): "Conflicts of Values". In: Bernard Williams: *Moral Luck: philosophical papers, 1973-1980*. Cambridge: Cambridge University Press 1981, 71-82.
- (1985): *Ethics and the Limits of Philosophy*. Abingdon: Routledge 2006.
- (1993): *Shame and Necessity*. Berkeley: University of California Press.
- (1995a): "Internal Reasons and the Obscurity of Blame". In: Bernard Williams: *Making sense of humanity and other philosophical papers, 1982-1993*. Cambridge: Cambridge University Press, 35-45.
- (1995b), "Replies". In: J. E. J Altham and R. Harrison (eds.): *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge: Cambridge University Press, 185-224.
- (1996a): "History, morality, and the test of reflection". In: Christine Korsgaard: *The Sources of Normativity*. Ed. by Onora O'Neill. Cambridge: Cambridge University Press, 210-218.
- (1996b): "Values, Reasons, and the Theory of Persuasion". In: Bernard Williams: *Philosophy as a Humanistic Discipline*. Selected, edited and with an introduction by A.W. Moore. Princeton/Oxford: Princeton University Press 2006, 109-118.
- (2001): "Postscript. Some Further Notes on Internal and External Reasons". In: Elijah Milgram (ed.): *Varieties of Practical Reasoning*. Cambridge: MIT Press, 91-97.
- (2002): *Truth and truthfulness: An essay in genealogy*. Princeton: Princeton University Press.

Wong, David (2006): "Moral Reasons: Internal and External". In: *Philosophy and Phenomenological Research* LXXII (3), 536-558.

Zahavi, Dan (2014): *Self & Other. Exploring Subjectivity, Empathy, and Shame*. Oxford: Oxford University Press.

Appendix

Abstract English

Bernard Williams famously argued that there are only internal reasons for action, that is, reasons which come up in sound deliberation on the basis of the agent's actual motivations. Two kinds of objections have been raised against this claim. Objectivists claim that there are external reasons which are independent of the motivations of actual agents. Intersubjectivists claim that there are external reasons which are not related to the agent's actual motivations but to the motivations of someone other than the agent. I argue that the objectivist objections to Williams are incorrect. Regarding intersubjectivist objections, I argue that current criticisms of Williams fail but that there might be an intersubjectivist concept of practical reasons that is satisfactory and important for accounting for certain considerations that come up in our practical deliberations. I put this treatment of Williams's internalism about reasons into the context of what I call a "theory of rational persuasion" and show what the consequences of the arguments regarding internalism and the objections to it are for the way we think about convincing others to do something.

Zusammenfassung

Bernard Williams vertrat die berühmte Ansicht, dass es nur interne Gründe für Handlungen gibt, d. h. Gründe, die in vernünftiger Überlegung ausgehend von den tatsächlichen Motivationen der Handelnden auftauchen. Gegen diese Behauptung sind zwei Arten von Einwänden erhoben worden. Objektivisten behaupten, dass es externe Gründe gibt, die unabhängig von den Motivationen der Handelnden sind. Intersubjektivisten behaupten, dass es externe Gründe gibt, die nicht mit den tatsächlichen Motivationen des Handelnden zusammenhängen, sondern mit den Motivationen von jemand anderem. Ich behaupte, dass die objektivistischen Einwände gegen Williams falsch sind. Was die intersubjektivistischen Einwände betrifft, so argumentiere ich, dass die derzeitige Kritik an Williams fehlschlägt, dass es aber einen intersubjektivistischen Begriff externer Gründe geben könnte, der konsistent ist und wichtig, um bestimmte praktische Überlegungen ausreichend erklären zu können. Ich stelle diese Behandlung von Williams Internalismus in Bezug auf Gründe in den Kontext dessen, was ich eine "Theorie der rationalen Überredung" nenne, und zeige, welche Konsequenzen die Argumente in Bezug auf den Internalismus und die Einwände dagegen für unser Nachdenken darüber haben, wie wir andere davon überzeugen, etwas zu tun.