THE UNIVERSITY *of* EDINBURGH

Edinburgh Law School

WHAT WE SAY WHEN WE CRIMINALISE – A METANORMATIVE INQUIRY

J.P. Fassnidge

Doctorate in Law

The University of Edinburgh

2023

ABSTRACT/LAY SUMMARY

Criminalisation has become a central topic for criminal law theorising. One idea that has drawn much scholarly attention is that criminalising a conduct involves, in some important way, the concept of wrongdoing. The conduct criminalised, it is usually said, are wrongs and this should play an important role in explaining both what is expressed through criminalising them and on how we should decide to criminalise said conduct. A recent trend in criminalisation theory has been to argue that the fact a criminalised conduct is, specifically, morally wrongful should be relevant to criminalisation in the way just described. If that were to be the case, then it becomes particularly important to have clarity as to what we mean by a conduct being morally wrongful and how we can go about determining that that is the case.

In parallel, there have been lively debates in philosophy about the nature and characteristics of morality, embodied primarily in the discipline of metaethics (sometimes also called meta-normativity). In these debates, accounts of what it means to say or to believe that something is morally wrongful vary greatly, to the point where holding different metaethical views entails completely different kinds of understandings of what it means for something to be 'morally wrong'. Being unclear about these differences risks talking past each other when trying to discuss moral wrongness. To date, these differences in metaethical stances have not been explored or taken up by criminalisation theorists, which leaves an important theoretical gap in criminalisation literature and theorising: if we mean different things by 'moral wrongness', we are not able to fully analyse and evaluate proposed theoretical frameworks for understanding criminalisation that use moral wrongness as a key concept in the framework. This thesis seeks to investigate

the connections between criminalisation and metaethics, under the assumption that the concept of wrongfulness is indeed to play any role in criminalisation, and show that criminalisation theorising would benefit from discussing these connections and from incorporating metaethical investigations into our normative theorising of criminalisation.

I argue that there are important differences between proposing that we are expressing, through criminalisation, that the criminalised conduct is morally wrongful, and proposing that we should use the supposed moral wrongfulness of the conduct in question as a deciding factor regarding its criminalisation. I explore these differences by making a distinction between different stages at which moral wrongfulness could potentially appear in the criminalisation process and its theorising. One way to approach criminalisation is to look at it as a process by which we use the law to make it the case that a conduct is, from now on, considered a criminal offence. In doing so, I argue that it is useful to view this action as a form of speech, by which we say something, through the act of criminalisation, about the conduct being criminalised – I call this the *declarative stage*. Another approach to theorising about criminalisation has been to view it as a decision-making process, by which a potential decision-maker is faced with the question of whether they should criminalise a conduct or not. In doing so, theorists can provide these decision-makers with normative frameworks under which they, supposedly, should make this decision and base it on sound normative principles – I call this the *deliberative stage*.

Once this distinction is made, I argue that both stages at which wrongfulness might be used in understanding criminalisation generate different possible meta-ethical questions about the nature and characteristics of that wrongfulness. I call these the conceptual, the ontological and the

epistemic questions, which I explore in detail in chapter 2.

The conceptual question asks about what is distinctive about the fact that the wrongfulness being invoked is specifically of a 'moral' kind. The ontological question asks about the characteristics of the existence (or lack thereof) of moral facts, properties and reasons. The epistemic question asks about access to, and the existence of, moral knowledge. By discussing different possible answers to these questions, this thesis intends to show that using 'wrongfulness' without giving a meta-ethical context for its use leaves many open questions that, if left unresolved, can entail very different understandings of what happens, and what should happen, when a conduct is criminalised.

Once the metaethical questions relevant to this thesis have been established, I then propose a framework for understanding each stage at which wrongfulness may appear within criminalisation, starting with the declarative stage in chapter 3. In it, I present an original framework of criminalisation as a speech-act, where different normative facts are included into the act of criminalisation. I argue that if wrongfulness is to play any role in what is conceptually involved in the act of criminalisation, then it must be as a kind of implicit assertion of the fact that the conduct being criminalised is morally wrongful. I then argue that, since this is an implicit assertion, a hearer of the speech-act needs to have enough contextual information in order to reasonably infer that the wrongness being asserted implicitly is of a moral kind. One way to argue that this is the case is that the declared normative facts included in the speech-act – an obligation not to do the conduct, a liability to punishment if one does, and a responsibility to answer for one's doing of the criminalised conduct – can be understood as the appropriate response to moral wrongdoing. I then show how this view

requires making important metaethical assumptions about moral wrongness, and I show these assumptions, and potential difficulties with them, through asking the conceptual, ontological and epistemic questions posed in chapter 2.

In chapter 4, I then move on to the deliberative stage, where I look at different normative theories of criminalisation which use the concept of wrongfulness in some important way. I explore some of the leading works of recent criminalisation theorists and how they use moral wrongfulness as a positive reason in favour of criminalising a conduct and/or as a negative constraint against criminalising non-wrongful conduct. I then apply the meta-ethical questions I have identified in order to construct potential answers that could represent the metaethical assumptions these theorists are making in their theorising. Particularly, different assumptions about the conceptual delimitation of moral wrongfulness, the ontological nature of the property 'moral wrongfulness' and the epistemic expectations one might impose on potential decision-makers as to how to gain epistemic access to moral wrongfulness can shift how some of those criminalisation theories work. I argue that in order to adequately use moral wrongfulness as a potential guiding principle (in its positive or negative form) for how we should decide whether to criminalise conduct, a theorist would benefit from providing a story as to what using that wrongfulness entails in meta-ethical terms, as this would avoid leaving theoretical gaps and would give clearer tools for potential decision-makers trying to apply these theories at the deliberative stage. Not doing so, in turn, can lead decision-makers to fill in the gaps and reach very different results, using the same deliberation process proposed by a theorist, because of making different metaethical assumptions.

Chapter 5 looks at a slightly different kind of theory, which presents itself

as an alternative to more orthodox 'legal moralist' theories and is critical of the use of moral wrongfulness as part of criminalisation theories, that has been dubbed by some the 'political turn' theories of criminalisation. These theories propose, in a nutshell, that instead of using 'moral' wrongness in our theorising of criminalisation, we should instead focus on a 'political' kind of normative fact in guiding decisions about criminalisation and criminal punishment. After analysing some examples of this kind of theorising, I show that they are dependent on important metaethical assumptions, specifically that moral facts and political facts are plausibly distinct. Then, I argue that there are two senses in which a political turn theorist can propose this distinction as intelligible: either as a subset of morality which is deemed 'political' (political moralist), or as a distinct normative domain, which is not the same as the moral (political realist). Both of these options, however, require the political turn theorist to choose between bullets to bite: either accept that political turn theories are less of an alternative to legal moralism, and theories that use moral wrongfulness, as they might have thought, or accept that they will need to commit to full-blooded political normativity, which will require them to do some heavy metaethical lifting.

ACKNOWLEDGMENTS

There are so many wonderful and generous friends and colleagues that I have met along the way in Old College, the PhD offices, conferences and seminars around the world and online, that have all played an important part in my academic growth and in getting to this point. Every single one has my gratitude, regardless of how large or small a part they have played in my career so far. I do wish to highlight a few colleagues and friends, old and new, that have had a particularly strong influence in both my work and my life, and dedicate this thesis and my gratitude to them.

To my supervisors, Dr. Andrew Cornford and Dr. Chloë Kennedy, who always pushed me to do my best work, while also supporting me when I was doubtful of my work and myself. Thank you for your patience, your understanding and your support throughout this whole journey. It has been an honour to work with both of you and I hope that I can live up to the trust that you have put in me and my ideas.

To the organising committee of the Virtual Criminal Law Group - Prof. Lindsay Farmer, Prof. Antony Duff and Prof. Sandra Marshall - who were generous enough to include me into what has turned out to be one of the most interesting academic groups I have ever been a part of. I have enjoyed every minute, every joke and every paper we got to discuss, and I look forward to many, many more. Also, thank you to all the members of the VCLG, who have seen some pieces of this thesis and listened to my ravings at every meeting we had. You are all a wonderful group of scholars and friends, and I hope that we can continue our conversations for many, many more years. A special mention for Prof. Alex Sarch, who very kindly read sections of this thesis and provided very useful insights into my work. I hope to be as generous and friendly a scholar as you have been to me.

To Dr. Martin Kelly, who was always interested in the work I was doing and would always encourage me to keep asking the big questions. Your friendship has been inspiring, and I hope to continue following each other's stories in the years to come.

To Dr. Jackson Allen, for being such a wonderful friend and for always helping me see the bigger picture for everything. We have been through thick and thin and pursued dreams together, including starting the Criminal Law Discussion Group at Edinburgh. The conversations we have had over the years have helped both of us to stay sane and remain enthusiastic about our work and our life choices. I love you like a brother and I'm lucky and delighted to have you, Lizzie and the incoming soybean in my life.

To Kajsa Dinesson, who will be Dr. Dinesson in a matter of months, for her enthusiasm and for always making me feel like I knew far more than I ever believed I did. Her unwavering support has always made me feel like my work and my thoughts were exciting and interesting, and I cannot thank her enough for all our conversations, coffees and lunches that helped form a wonderful friendship. I look forward to further adventures and flat whites.

To Dr. Grant Barclay for sharing his joy and his friendship with me. It has been a pleasure to talk about criminal law, and more importantly music and videogames, with someone both as knowledgeable and funny as him. I hope to continue our rants on progressive metal and RPGs for many years to come.

To Gabriel Muñoz, for being like an older brother and a dear friend since the first few weeks I came back to the UK and began my Edinburgh adventure. Here's to many, many more years of drams, football and rants on politics and law. Also, to Stefan Domancic, who always believed in me and my dreams: here's the result, brothah. Thank you and onwards!

To my wife Lore, who supported me every step of the way in this journey, and always did her best to be understanding of the bizarre process that doing a PhD entails. None of this would have been possible without her help, her patience, and her bravery in facing whatever came our way. I am lucky to have her as my partner and my friend. I love you with all my heart.

To my daughter Ellinore, for whom all these efforts were done. You have shown me what really matters in life and not a day goes by when I don't smile because of you. I love you, Ellie, with everything that I am.

Finally, to my father, Chris Fassnidge, whose life inspired me to pursue this academic journey in the first place. I wish he was here to see how his example lit the way for the kind of scholar and thinker I wanted to be. I will forever miss you and carry your memory with me.

This doctoral research was kindly funded by CONICYT (Comisión Nacional de Investigación Científica y Tecnológica/Chilean National Commission for Scientific and Technological Research), now ANID (Asociación Nacional de Investigación y Desarrollo/National Agency for Research and Development) "Becas de Doctorado en el Extranjero Becas Chile, Convocatoria 2018", Grant No. 72190024.

J.P. Fassnidge
Edinburgh, 2023

Table of Contents

CHAPTER 1 – CRIMINALISATION, WRONGFULNESS AND METAETHICS

Criminalisation, broadly understood as the process by which conduct becomes a criminal offence within a particular jurisdiction, has become a central topic of discussion in criminal law theory. In particular, an important part of the literature on criminalisation theory has focused on the role that moral wrongfulness plays in both explaining what happens through criminalisation, as well as for normative theories on how criminalisation should be decided.[1] Parallel to these debates around criminalisation, metaethics – the area of philosophy that "aims to explain how actual ethical thought and talk—and what (if anything) that thought and talk is distinctively about—fits into reality"[2] – has developed a rich discussion around what is happening when someone uses moral talk or thought. By trying to analyse and propose how moral talk and thought actually works and its characteristics from different philosophical points of view, the adoption of specific metaethical commitments can lead to different understandings — indeed, different meanings — of what is happening when we *say* and *think* that something is morally wrong.

[1] A few recent examples, which shall be discussed in greater detail throughout the thesis are Michael S Moore, *Placing Blame - A General Theory of the Criminal Law* (Oxford University Press 1997); Douglas Husak, *Overcriminalization: The Limits of the Law* (Oxford University Publlishing 2008); RA Duff, *Answering for Crime - Responsibility and Liability in the Criminal Law* (Hart Publishing 2009); AP Simester and Andreas Von Hirsch, *Crimes, Harms and Wrongs - On the Principles of Criminalisation* (Hart Publishing 2011); Victor Tadros, *Wrongs and Crimes* (Oxford University Press 2016); RA Duff, *The Realm of Criminal Law* (Oxford University Press 2018); Thomas Søbirk Petersen, *Why Criminalize?: New Perspectives on Normative Principles of Criminalization*, vol 134 (Springer International Publishing 2020) <http://link.springer.com/10.1007/978-3-030-34690-4> accessed 16 May 2023.

[2] Tristram McPherson and David Plunkett, 'The Nature and Explanatory Ambitions of Metaethics' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2018) 3.

This thesis seeks to look at current theoretical understandings of criminalisation through the lens of metaethics. It is an effort to show how, by using the theoretical tools provided by metaethics, it is possible to expand our knowledge and understanding of criminalisation theories that use, or could use, moral wrongfulness in some important way. It is not, however, a defence of any particular metaethical view, but rather an effort to show how different metaethical commitments can affect our theorising of criminalisation, and how informing our normative theorising with metaethical investigations is a better route forward than not doing so.[3] This theoretical approach has not been attempted before in criminalisation theory, and this thesis seeks to contribute some basic tools from metaethics in order to look at criminalisation from a different and novel perspective, while also making an original contribution by pointing out a gap in our criminalisation theorising, regarding the implications of different metaethical stances to our understanding and analysis of criminalisation.

The importance of taking this approach is that if we think that moral wrongfulness is in some way involved in the process of criminalisation – and, as we shall see, many people do – then I argue it is important to address the differences that holding metaethical stances can have on our use of moral wrongfulness, because if we are not talking about the same things when talking about 'wrongfulness', as what is being stated or used to consider justifying criminalisation, then we risk losing any meaningful understanding or critical analysis of what is being discussed. We risk *meaning completely different things*. We might end up just talking past each other.

In this chapter, I will introduce the initial theoretical framework which will

[3] This effort is in large part inspired by the ideas proposed in Tristram McPherson, 'Unifying Moral Methodology' (2012) 93 Pacific Philosophical Quarterly 523.

be necessary for the rest of the analysis carried out in the thesis. The proposed framework looks to offer an original view on how one could go about analysing criminalisation from a theoretical perspective, and in doing so I show how the concept of wrongfulness can appear within those perspectives. In this way, I seek to make clear what the theoretical targets of the thesis are, as well as provide answers for connected, but importantly different, questions about criminalisation. These are, first, how can moral wrongfulness play a role (if at all) in theories of what is being *said* when criminalising a conduct; second, how does moral wrongfulness play a role (if at all) in the deliberation processes for criminalisation proposed by normative criminalisation theorists; and third, how do different metaethical commitments alter the theoretical understanding of criminalisation, both for what is being said and for how decision-makers should deliberate on what conduct to criminalise.

This thesis argues that when trying to understand criminalisation, one can take different theoretical approaches. I focus on two possibilities: understanding criminalisation as a form of communication, and understanding criminalisation as a decision-making process. Though both approaches concern the process by which a type of conduct becomes a criminal offence, the first approach tries to analyse what is potentially being communicated by the act of criminalising a conduct, whereas the second approach focuses on the prior process by which a potential decision-maker is supposed to decide whether a conduct should be criminalised or not. Both approaches, however, can complement each other in the sense that one can propose a certain understanding of the decision-making process, where the specific target of the decision is whether we should communicate whatever is being communicated when a conduct is criminalised.

Once these approaches have been established, I will then look at each approach in more detail and how wrongfulness can play a role at each stage. Then, I discuss why it might make sense to think that wrongfulness has a role to play in the understanding of criminalisation. I argue that if wrongfulness has a role to play – and there may well be good reasons to think so – then it becomes necessary to analyse and examine the metaethical implications of using wrongfulness to theorise about criminalisation. To do this, I present a set of questions – which I call the conceptual, ontological, and epistemic questions – whose answers will provide us with the theoretical materials necessary to understand the metaethical implications of using wrongfulness for the purposes set out above. I then explain why it is important to make these metaethical inquiries, and argue that not doing so risks talking past each other when trying to understand criminalisation and explain its structures through using wrongfulness. Finally, I conclude the chapter by providing a roadmap for the rest of the thesis.

1.1. Criminalisation as Communication and as Decision-Making through Wrongfulness

Let us begin by looking at criminalisation through the lens of communication. As I said earlier, one approach to understanding criminalisation is to focus on the act of criminalisation itself – the act(s) by which a type of conduct becomes a criminal offence within a legal system. Under this kind of approach, the criminal law becomes, among many other things, a tool for communication. Criminalisation is, in itself, a form of said communication. It says something about the conduct being criminalised and we might expect people within a particular jurisdiction to be able to understand what is being said by criminalising the conduct. Even if a visitor in said jurisdiction were to be made aware that a certain conduct is

criminalised, it is not a farfetched expectation that they too would understand what it means for that conduct to be criminalised. The understanding would, one would expect, go further than observing that a conduct is included into the criminal law, or that a particular conduct falls within the legal category 'criminal offence'.[4] That is, formally defining a conduct as a criminal offence – usually done by the legislature but it does not necessarily have to be them – seems to communicate certain things to the people within a jurisdiction. This is because the criminal law – and the law more generally – deals in actions and conduct,[5] and where there are actions there is potential for meaning and the communication of said meaning. In this sense, as Van Hoecke has pointed out, "human action implies interpersonal relations and, thus, communication. As a consequence, if law offers a framework for human action, it also offers a framework for human communication."[6] I will call this approach of theorising on criminalisation as a theory of the *declarative stage* of criminalisation.

One thing that we might be saying through the act of criminalising is that the conduct being criminalised is, in some important way, *wrongful*. That is, in making a conduct count as a criminal offence, we are communicating that that conduct is, in some sense, the wrong thing to do. There are various conceptual distinctions that I will explore further below regarding the nature of wrongdoing, but the first step in the discussion is to simply understand the possibility of linking the act of criminalisation to the conveying of wrongfulness of the criminalised conduct.

[4] Of course, these things are also included into the broader meaning of criminalising a conduct, which I am not denying. The point, rather, is to say that the understanding that can be reached by knowing that a conduct is criminalised goes beyond only the aforementioned facts.
[5] We can use 'action' and 'conduct' interchangeably, but I simply add the term 'conduct' here as a slightly distinct concept, in the sense that a 'conduct' can be a succession of 'actions' that are understood as one distinct 'conduct'.
[6] Mark Van Hoecke, *Law as Communication* (Hart Publishing 2002) 7.

Some might want to go further even and say that what is being said is that the conduct is *morally wrongful*. I will explore the meaning of this distinction further throughout the thesis, but for now it is important to understand that there are two potential senses in which one can read 'wrongness' or 'wrongfulness' in the context of this debate, and more broadly in any kind of normative debate. The first sense is a narrow understanding of 'wrong' as referring to that which is morally wrong, i.e. that which is wrong according to morality. The second sense is a broader understanding of 'wrong' as a more general normative concept – something which is wrong with regards to some kind of normative standard, as in the wrong move in a game, the wrong choice for obtaining a particular goal, the wrong gesture or action according to etiquette, and so on. When I speak of 'wrongs' or 'wrongness' *simpliciter*, I will be referring to this latter sense of wrongs, and I will explicitly say 'moral wrongs' or 'moral wrongness' when I mean the narrower sense of wrongfulness.

A different but related approach is to try and understand criminalisation not just from looking at the act of criminalisation itself, but instead at the prior stage of *deciding* that a conduct *should* or *should not* be criminalised. This kind of approach is, as one might imagine, much more normative in nature, in the sense that it is not focused exclusively on what criminalisation does in practice, but rather on how that process for the decision to criminalise a conduct should be carried out in the first place. It is trying to answer a normative question – whether a type of conduct should be criminalised – instead of a descriptive one – what happens when a conduct is criminalised. This divide, however, is only suggested for analytic purposes, as the two approaches inform each other: by understanding what happens when a conduct is criminalised, we can better inform our theories on when that should happen. I will call this approach a theory for the *deliberative stage* of

criminalisation.

Importantly, this differentiation of approaches is proposed to better appreciate how wrongfulness can be used within criminalisation. It can play a role as part of the explanation of criminalisation as an act of communication – we say something about the wrongfulness of the conduct criminalised – or it can play a role in how we should decide to criminalise a conduct – wrongfulness helps us explain how we should decide whether to criminalise a conduct or not. If this is the case, then it becomes particularly important to explore both how those uses of wrongfulness work and what using the concept of wrongfulness entails.

This distinction will allow for a more nuanced examination of criminalisation theories that have been proposed in recent literature, where the concept of moral wrongfulness has become increasingly important as a part of what is involved in criminalisation as a whole. Some argue that when we criminalise a conduct, we are declaring the conduct to be morally wrongful;[7] some believe that the fact that a conduct is morally wrongful counts as a positive reason in favour of criminalising that conduct;[8] some argue that moral wrongfulness is a constraint on criminalisation, and that if a conduct is not morally wrongful then one lacks a reason to criminalise it;[9] some go as far as to say that the criminal law 'speaks with a moral voice'.[10] All of these views have in common the idea that wrongfulness, understood particularly as *moral* wrongfulness, has a significant role to play in criminalisation. The thesis will argue, however, that different metaethical understandings of what moral wrongfulness actually is and what it entails can

[7] Duff, *Answering for Crime - Responsibility and Liability in the Criminal Law* (n 1).
[8] Moore, *Placing Blame - A General Theory of the Criminal Law* (n 1).
[9] Husak, *Overcriminalization: The Limits of the Law* (n 1).
[10] Simester and Von Hirsch (n 1).

lead to important differences in our understanding of what is being said through a criminalisation claim, and of how a potential decision-maker is supposed to decide whether to utter a criminalisation claim or not. These differences become apparent through a metaethical analysis of both stages and the implications of making different metaethical assumptions for the use of moral wrongfulness as both an instance of moral talk at the declarative stage and as a moral property to identify at the deliberative stage.

1.2. Stages of Analysis for Criminalisation and Wrongfulness: Declarative and Deliberative stages

The formal inclusion of a particular conduct into the criminal law as a criminal offence is done by some kind of institutional act – often, by an act of Parliament or the legislature. My focus will be on how we theorise about this formal act, but I propose to place this formal inclusion against the backdrop of a larger decision-making process. The act by which a conduct is criminalised does not happen naturally or spontaneously – it is the result of a series of actions and decisions that culminate in the final product of a criminalising act. My purpose in this thesis is not to look at every action that takes place during that process leading to criminalising a conduct, but instead to set up a conceptual framework that makes sense for a better understanding of how criminalisation works and, importantly, how we theorise about it.

First, however, I need to explain what I mean by 'formal' criminalisation. Nicola Lacey introduced the distinction between 'formal' and 'substantive' criminalisation,[11] both for understanding criminalisation as an outcome and

[11] Nicola Lacey, 'Historicising Criminalisation: Conceptual and Empirical Issues' (2009) 72 Modern Law Review 936.

as a social practice. Formal criminalisation can be understood as "in the books" and substantive criminalisation as "in action", where the former manifests through legislation, judicial decisions and international treaties, and the latter manifests in the actual implementation of formal norms in the actions of agents within the criminal justice system.[12] My focus in this thesis will be mostly on the formal side of Lacey's spectrum – my claims are made regarding the process by which an action gets to be officially categorised as a criminal offence in the first place.[13] This does not mean that the interest is only forward-looking towards potentially new offences – one can apply the same analysis I will propose to already existing offences retrospectively, as in asking ourselves how and why a criminalised conduct got to be categorised as such when it did. My interest, however, is mostly normative rather than descriptive. I will not be looking at the actual ways in which a conduct gets to become an offence – the focus is not, for example, on the legislative process itself and the factors that may, as a matter of practice, affect how this is done. Rather, my interest is at a more conceptual level, in the sense that the framework I will propose seeks to understand, in the abstract, how formal criminalisation works. This does not mean that my interest is purely in ideal or in normative theory – the focus will not be solely on how a criminalisation process *should* be nor how it ought to be in the ideal, even conceptually. Instead, my goal is to set up a framework to look at the formal process of criminalisation from two theoretical perspectives: how we theorise about what happens when a conduct is criminalised (i.e. the end-product of the process) and how we think about what should be taken into account in order to get to that end-stage.

[12] ibid 943.

[13] There are, of course, very interesting questions to be made about substantive criminalisation and how metaethics might affect criminalisation 'in practice', but they unfortunately escape the scope of this thesis.

I have also roughly defined criminalisation as the process by which a conduct is classified as a criminal offence. This is a rough definition because in practice criminalisation, as we shall see further below, does more than just define conducts as criminal offences. In particular, the effects of what I will be analysing throughout this thesis – the formal act of criminalising a conduct – that occur *after* the portion of the criminalisation process I will be analysing, are both part of the broader process of criminalisation and are important considerations for understanding what should happen *before* this formal act takes place. This is particularly true with regards to what many scholars see as (at least one of) the defining characteristic(s) of the criminal law – its condemnatory and punishing functions. Tadros explains the line of thought that generally links the use of wrongfulness to how punishment takes a part in criminalisation as follows: "two of the central functions of criminal law are condemnation and punishment of offenders. But condemnation and punishment are justified only if the person has done wrong. Therefore, it is wrong to criminalise conduct that is not wrong."[14] Regardless of whether one agrees with this line of argument – and Tadros does not– the point I am trying to make here is that the reason this kind of concern becomes relevant to the debate around criminalisation, as Cornford rightly points out, requires committing to two claims: a conceptual claim about criminalisation, where "to criminalise conduct is to warrant the condemnation and punishment of that conduct";[15] and a normative claim with regards to when it is permissible to condemn and punish, where "one may condemn and punish others only for their wrongful conduct."[16] Regardless of whether one agrees with condemnation and punishment playing such an important role for criminalisation, what is true is that criminalising a conduct *does* currently

[14] Victor Tadros, 'Wrongness and Criminalization' in Andrei Marmor (ed), *The Routledge Companion to Philosophy of Law* (Routledge 2012) 165.
[15] Andrew Cornford, 'Rethinking the Wrongness Constraint on Criminalisation' (2017) 36 Law and Philosophy 615, 617.
[16] ibid 622.

create a liability to punishment or to consequences that can be understood as punishment, and so that fact does need to be taken into account in our understanding of how criminalisation works. Hence, at least in part, why I am limiting the scope of the analysis to what I am calling formal criminalisation.

I propose, then, that the best way to understand formal criminalisation is to split the analysis into two stages, namely the formal act of criminalising a conduct itself and the decision-making process by which we end up with the act of criminalising. As already mentioned, I will call these stages of analysis the *declarative* and the *deliberative* stages respectively. In the declarative stage, the focus of analysis is the conceptual understanding of what it actually is to formally criminalise a conduct – the target is the formal act of criminalisation in itself, which is the end product of the criminalisation process. In the deliberative stage, the focus is more on the ways in which a theorist proposes that a decision-maker should decide whether to perform the formal act of criminalisation, and in particular this is where normative reasons for or against criminalisation can be proposed.

1.3. Is wrongfulness necessary to understand criminalisation?

For this thesis, I will be analysing the idea of wrongfulness broadly and how it can have a place within criminalisation. As mentioned earlier, I make a distinction between a broad and a narrow sense of 'wrongfulness' or 'wrongs', and I will be looking at whether both kinds of wrongness (i.e. wrongness *simpliciter* and moral wrongfulness specifically) can or should have a role to play in our understanding of criminalisation. That being said, one potential problem we could have has to do with this entire approach to criminalisation in the first place – why think that wrongfulness (broad or narrow) has any role to play in criminalisation? In other words, why must our theorising about

criminal law, including criminalisation, deal with wrongs at all?

I am willing to grant the point of the question with regards to wrongfulness in a narrow sense. In fact, it is a question that underlies a lot of my analysis — one needs to explain why morality is the appropriate or relevant normative domain for criminalisation, and to do so two prior questions need to be addressed: what does 'morality' mean in this context (which I will address in chapter 2) and why is it relevant to the criminal law. One possible answer, as signalled above when discussing the definition of criminalisation, is to point to the fact that condemnation and punishment are part of the central functions of criminal law, and that the only way to justify these practices is by using moral wrongfulness. Though I will not attempt here to answer the question of function, I will look at why one might think the moral domain is a relevant normative domain for criminalisation more closely in chapter 2, by looking at how one might answer the conceptual question posed there. Note, however, that this is a slightly different question to the one I posed earlier — it is not on why we might think the criminal law deals with wrongs, but specifically it is about the relevance of *morality* as the right normative domain from which to obtain the wrongs that criminalisation purportedly deals with.

With regards to wrongfulness more broadly, however, the question becomes more problematic. That is, believing that any kind of wrongfulness has no role to play within criminalisation entails that, for example, criminal offences do not include a wrong *in any sense*. This means that the conduct or action included within a criminal offence would *not* be able to be considered a kind of shortcoming with regards to some kind of normative standard — it would not be an action that can elicit appropriately a response of 'that's not the right thing to do' in *any* sense. In other words, it would require us to accept that including a conduct into a criminal offence is not, in any important

sense, normative in the first place – creating a criminal offence would not involve stating a norm *about* the conduct in question.[17] This is because, as Gerritsen puts it, norms are "standards, rules, principles or conventions that involve correctness conditions, rather than statistical normalities. Norms involve correctness conditions because they categorise our behaviour and attitudes as correct or incorrect, depending on whether they are in accordance with the norm in question."[18] Accepting that criminal offences do not include the statement of some kind of norm about the conduct being criminalised seems counter-intuitive, as it would lead to accepting this kind of *modus tollens* argument:

(1) If criminal offences include conduct that are wrongs, then criminal offences include a statement of norms (i.e. standards including correctness conditions for behaviour) about the conduct.

(2) It is not the case that criminal offences include a statement of norms about the conduct.

Therefore, (from (1) and (2)),

(3) It is not the case that criminal offences include conduct that are wrongs.

Premise (1) is a reflection of what I have explained above: the presence of a concept of wrongness necessitates there being a norm involved for there to be something wrong in the first place. Premise (2) does not, however, seem plausibly true. This is because norms establish favouring or disfavouring relations – based on correctness conditions – between an action (in this case, the criminalised conduct) and a fact that counts in favour or against the action.

[17] By 'stating' here I leave open the possibility of it meaning either reporting the pre-existence of a norm about the conduct, or creating a norm about the conduct in the act of stating it.

[18] Eline Gerritsen, 'Demistifying Normativity - Morality, Error Theory and the Authority of Norms' (PhD Thesis, Groningen/St Andrews/Stirling 2022) 32 <https://research-repository.st-andrews.ac.uk/handle/10023/27003>.

If this relation seems like a good explanation for what is going on with a criminal offence – that is, making a conduct count as an offence is best explained as establishing a normative relation between the conduct and a fact that favours or disfavours its performance – then criminal offences must be (or include) some kind of norm, and therefore at least some form of wrongness needs to be present in the fact that it is being criminalised (albeit, not necessarily of a moral kind). So, at the very least, if a criminal offence can be seen as including some kind of norm about the conduct in question, then at least some form of wrongness will necessarily be involved – wrongness in the sense of incorrectness with regards to being in accordance with the norm in question. And, on the contrary, if we were to argue that no form of wrongfulness is involved, then criminal offences could not be seen as including norms of some kind about the criminalised conduct, moral or otherwise. I will explore these ideas in more detail in chapter 3, where I discuss the declarative stage of criminalisation, but the point for now is to show that, from a theoretical perspective, it does make sense to think of wrongfulness in a broad sense as an appropriate conceptual tool to understand how criminalisation manifests.

One way of seeing this normative relationship between an action and a fact that counts in favour of that action would be to reduce it to a kind of norm of categorisation – one could correctly or incorrectly classify a particular conduct as the criminal offence being stated. And this, undoubtedly, is an important part of the process when working with crimes – it is precisely why the elements of an offence are so important to establish when looking at a potential case of offending, in order to correctly impute a conduct to a particular offence. However, reducing offences exclusively to a labelling exercise seems incorrect, in large part because there is a lot more happening through criminalisation and the consequences thereof, as I shall explore in

more detail in chapter 3 when looking at the declarative stage. For now, however, we can already say that, at the very least, some form of punishment is *also* being established in the act of criminalising conduct, so it cannot be the case that offences only partake in a labelling exercise.

To summarise: though there are legitimate questions to be asked as to why *moral* wrongfulness has any role to play within criminalisation, the analysis I will propose in this thesis is based on a conditional statement: *if* we were to believe that moral wrongfulness *has* some role to play in criminalisation, then it is important to consider the metaethical issues I will raise throughout the thesis. Of course, the metaethical issues around moral wrongfulness can also be extended to wrongfulness broadly understood, so the framework I intend to provide can also be helpful for those who are sceptical of using moral wrongfulness within criminalisation *if* they are willing to accept the idea that wrongfulness broadly understood does have a role in criminalisation.

1.4. Metaethical inquiries for criminalisation

As already stated, I will pose some questions about the use of moral wrongfulness within the context of criminalisation and its theorising, which can be answered by using tools from metaethics.[19] These questions can broadly be categorised into three kinds of general inquiry. The first I call the *conceptual* question, which relates to identifying and defining what is distinctive about *moral* wrongfulness. In essence, the conceptual question targets the definition and delimitation of what counts as morality, which as we

[19] There is also the term 'meta-normative' theory, which I will sometimes use interchangeably with metaethics throughout the thesis. The only potential difference between the two terms is that 'meta-normative' is applying the same kind of investigations as metaethics (which usually focuses on morality), but to all sorts of potential normative domains, such as epistemology, rationality, prudence, and so on.

shall see involves questions about the characteristics of morality as a normative system and its relationship to other normative contexts. The second question I call the *ontological* question, which focuses on trying to understand the nature of moral wrongfulness – once we have defined what makes wrongfulness moral, we can then ask ourselves about the characteristics of its existence as a fact or a property. This will involve metaphysical and ontological questions about moral wrongfulness and we shall see that there will be different kinds of issues around moral wrongfulness as it features in moral talk and as it features in moral thought. Finally, the third question I call the *epistemic* question, which looks at the epistemological issues around our acquiring knowledge about moral wrongfulness – the focus is on how we gain epistemic access to moral wrongfulness.

In turn, these broad categories of questions will manifest in more specific issues around each stage of analysis of criminalisation which I have identified above. At the declarative stage, the conceptual question will seek to understand how 'moral wrong' is being used as a form of speech, and what makes it the case that it is specifically a case of 'moral talk'. The ontological question will look at what things in the world (if at all) the speech-act is trying to talk about, and where does that use get its meaning from. Finally, the epistemic question will allow to analyse issues around gaining access to the knowledge about the meaning of the speech-act, and how that knowledge is obtained.

Contrastingly, the conceptual question at the deliberative stage focuses on what senses of 'moral wrong' are being used in theorising decision-making for criminalisation. The ontological question will give us tools to understand how theorists think the moral wrongness included in their theory

manifests as a *property* of criminalised conduct. Finally, the epistemic question at the deliberative stage will allow us to ask the important question on how a purported decision-maker is supposed to *find* that property of moral wrongness, and which kind of knowledge-acquiring process is best to do so.

Notice that these three categories of questions are distinct but also, at least to a certain degree, inter-dependent — how we answer the conceptual question affects how we answer the ontological question, and how we answer this latter question affects how we answer the epistemic question. By knowing the boundaries of what we are to analyse (i.e. answering the conceptual question) we will know what facts and properties we need to identify in order to answer the ontological question, and once we are clear on what facts and properties are relevant we can ask ourselves how we gain epistemic access to them in order to answer the epistemic question. There are, of course, further interesting questions to be asked from the perspective of metaethics about moral wrongfulness, like the phenomenology of moral qualities and how we get to experience them as entities, or the psychological questions around the motivational state of someone making a moral judgment and the connections between making a moral judgment and the motivation to do as the judgment says.[20] Though perhaps some of the issues raised by these questions will be relevant to the analysis carried out in this thesis, I will circumscribe what may be relevant into the three broad categories I established earlier, both for purposes of ease of understanding and to not cast too wide a net of metaethical issues being raised here.

[20] See Alexander Miller, *Contemporary Metaethics - An Introduction* (2nd edn, Polity Press 2013) 2–3.

1.5. Why do metaethical inquiries about criminalisation?

One question that needs to be answered at this stage is on the importance of using metaethical investigations for normative theorising in the first place. After all, there has been an assumption in a part of philosophical literature that metaethics is irrelevant or unnecessary for normative theorising.[21] The main strategy to argue this has been to state that it is plausible to think that two people might agree on what the correct normative theory is, but disagree upon their metaethical theories.[22] However, this thesis adopts a contrary view, under the understanding that, as Kagan put it, "doing normative ethics requires having views about metaethical issues."[23] In this sense, I agree with McPherson's 'unity thesis', whereby "the best way to pursue the goal of normative theorizing involves *integrating* our normative theorizing with systematic metaethical inquiry. In short, this involves pursuing the answers to metaethical and normative questions together."[24] My proposal in this thesis is precisely to integrate metaethical considerations into our normative criminalisation theorising, because I argue that doing so will improve our criminalisation theories.

A good example for how metaethics can better inform our normative theorising is the different kinds of epistemic methodologies one can use in deciding within a normative framework, such as the deliberation processes for criminalisation proposed by theorists. The idea is that in exploring the

[21] For a particularly strong version of this view, see Ronald Dworkin, 'Objectivity and Truth: You'd Better Believe It' (1996) 25 Philosophy and Public Affairs 87; Tristram McPherson, 'Metaethics and the Autonomy of Morality' (2008) 8.
[22] A historical example of this view can be found in William K Frankena, 'Main Trends in Recent Philosophy: Moral Philosophy at Mid-Century' (1951) 69 Philosophical Review 44.
[23] Shelly Kagan, *Normative Ethics* (Westview Press 1998) 5.
[24] McPherson, 'Unifying Moral Methodology' (n 3) 525–526. Emphasis in the original.

different epistemic methodologies for finding moral facts, one gets to work one's way backwards to figure out the metaethics that informs said methodology. Here is an example based on the discussion presented by McPherson.[25] Suppose that one thinks moral wrongness is found by carefully consulting one's intuitive reactions to cases (moral intuitionism). What kind of metaethics justifies or vindicates using intuition as the preferred method for finding moral facts? And if one has a different kind of metaethics, say naturalistic realism about moral facts,[26] then said view is going to be inconsistent with an intuitionist methodology – it would be like trying to understand the nature of water by appealing to one's intuitions. So, if we were still to think that intuition *can* actually inform our epistemic methods in understanding something like moral wrongness, then we need to provide a metaethical framework that is consistent with using intuition to find moral facts. Additionally, we will need to explain why that metaethics makes more sense or is preferrable to the naturalistic realism that is incompatible with intuitionism. This will be possible if we use the tools of metaethics, and find our answers to the metaethical questions I have posed in this thesis, when doing normative theorising. That is, the methodological controversies that can arise from our different normative theories and their justification can be best addressed by using metaethics and its investigations into the nature of moral thought and talk to inform our differences. As McPherson puts it, "we all need principled ways of resolving the methodological controversy that confronts our project, and metaethics is one promising source for discovering such principled resolution."[27]

McPherson also gives the example of burgeoning research in moral

[25] ibid 539–540.
[26] I will explore this view in more detail in chapter 2, but in a nutshell it is the view that there are true moral facts in reality, and that they can be found as 'natural' facts, in the same ways as one might find facts about electricity or gravity.
[27] McPherson, 'Unifying Moral Methodology' (n 3) 540.

psychology and how some have suggested that, for example, "our tendencies to judge that there is a moral asymmetry between doing harm and allowing harm to be done, or between intending harm and merely foreseeing it as a side-effect of one's action, can be explained away as results of allegedly morally arbitrary psychological processes", and that it is vital to include metaethics into the debate about the normative significance of these facts, in view of the fact that "accepting different accounts of the nature of moral concepts or moral properties may well lead to very different evaluations of the significance of these psychological processes, and the research that explores them."[28] And it is this kind of influence on normative theories which can also inform our theorising of criminalisation more generally, and of the deliberation process for it more specifically. The way in which one understands the nature of moral wrongness will lead to different evaluations of the epistemic methods by which that wrongness is identified and used within a deliberation process.

Another important aspect to consider in favour of being clear about the moral epistemology being assumed for the deliberation process, is how potential disagreements with the judgement of the existence or absence of a moral status by the decision-maker can be dealt with. In having clarity on how a judgement has been formed from an epistemic point of view, and in knowing whether there is a cognitive state like a belief in a moral fact in the first place, we can evaluate whether our disagreement with the judgement of the decision-maker (or with the deliberation process the potential decision-maker would be following) is one based on some kind of epistemic mistake. That is, we can now check whether something in the epistemic method being used (or being proposed as part of the deliberation) has gone wrong, or whether we think that the method itself is incorrect for the kind of inquiry

[28] ibid.

required for the decision on criminalisation. So, for example, we can check which pieces of evidence (if any) have been used to justify the belief in the presence of moral wrongness, and we can check whether said evidence is what we would consider appropriate for justifying that belief. This kind of revising, in turn, requires clarity not only for the methodology applied in seeking the evidence, but also the metaethical assumption that something like moral wrongness is the kind of thing that is susceptible to a cognitive process; that moral talk is truth-apt and that we can form true or false beliefs about moral properties.

These kinds of substantial differences can also be extrapolated to the rest of our theorising about criminalisation, when using moral wrongness. Our metaethical assumptions on how moral talk and thought works will end up shaping not only our discussions of what we say when a conduct is criminalised and how we decide to say this, but they will also entail many further questions that, depending on how we answer them, can lead to very different understandings of what the criminal law is doing when criminalising, not to mention how we understand if that is what it should be doing. To give an example, suppose we have moral realist assumptions about moral talk – moral talk tracks mind-independent facts and can be either true or false. If this is right, then we need to deal with – to only give a few examples – the following questions about moral talk in the context of criminalisation: how do we find these moral facts? Are these facts natural or non-natural? Do we require a particular skill, knowledge or perceptive ability to find them? If so, do we expect those who decide about criminalisation to have these skills, knowledge or abilities? Do we expect everyone to have these abilities? What happens if we have a disagreement on these facts? There are also further assumptions entailed in moral realism that affect moral talk, like the fact that if moral properties are mind-independent, then our moral talk about moral

properties will mostly be a set of assertions about the world, and our words will be trying to track those properties in the world.

In a way, it is similar to the way in which maths teachers would ask us in school to "show our work". We might end up reaching the same result, but through different methods. But importantly, if we have arrived at different results it might not be that one of us is mistaken – we might have just used different processes and methods to get to our results. Additionally, by looking at each other's methods, we can more accurately determine if something has gone wrong and how. However, if we simply do not talk about the methods by which those results are obtained, we might never be able to accurately understand each other, and we might even think that our differences are intractable or based on something else (like, in the case of moral wrongs, ideology or irrationality). Instead, it is more fruitful to our criminalisation debates to be clear as to what exactly we mean by a conduct being wrong (morally or otherwise), what we think makes it the case in our world that said conduct is wrong, and how we expect decision-makers to figure this out in their deliberations.

Finally, a problem that can be dealt with by including metaethical investigations in our criminalisation theorising has to do with disagreeing on *what we are doing* when judging something to be morally wrongful. This problem may seem subtle at first glance, but it runs much deeper when looked at closely because it can lead to those in disagreement talking past each other. If we are *doing different things* when making claims about wrongfulness, then we are not even talking about the same things, so we cannot genuinely disagree and instead are simply not coming to any real kind of understanding. More importantly, if we are not talking about the same thing when talking about 'wrongfulness' as what is being communicated

through criminalisation, or as a reason that can or cannot justify criminalising a conduct, then we cannot really have any meaningful understanding or critical analysis of what is being discussed, because *we mean completely different things*. We are just talking past each other. Therefore, it is best if we are clear in our metaethical commitments, and the metaethical issues that are involved in normative theorising, so as to avoid this talking past each other, as well as to have a more complete theoretical understanding of what criminalisation involves.

1.6. Roadmap for the Thesis

Now that the framework for analysis is clearer, I will set out the roadmap for the rest of the thesis. Chapter 2 will provide the necessary knowledge base to understand the metaethical issues I will be discussing throughout the thesis. In it, I will lay out the conceptual, ontological and epistemic questions, summarising some of the basic literature in metaethics that gives content to these questions.

Next, in chapter 3, I will look at the declarative stage of criminalisation. To better understand what might be communicated through the act of criminalising a conduct and how that communication could happen, I will make an original contribution to the field by proposing a model for what is being said when a conduct is criminalised, based on speech-act theory. I will argue that when conduct is criminalised, it can be best understood as a kind of speech-act, by which the speaker (in most cases, the legislature) both asserts a property of the criminalised conduct and declares into existence new normative facts regarding that conduct, and in doing so makes it the case that the conduct is hereby a criminal offence. I will then argue that if moral wrongfulness is to have any role to play in this speech-act, then it will be as

the asserted property of the criminalised conduct. Importantly, however, this assertion is done implicitly through the speech-act, and therefore needs to be inferred by the hearer of the act through context. I will then use the metaethical questions posed in chapter 2 to show how different metaethical stances change our understanding of how this implicit assertion of moral wrongfulness would work.

Chapter 4 focuses on the deliberative stage of criminalisation. In it, I look at some criminalisation theories that have made prominent use of the concept of moral wrongness for the normative theorising on criminalisation. For each theorist, I then reconstruct what their proposed deliberation process for a criminalisation decision-maker looks like, and then construct from their work how they might answer the metaethical questions I have posed in this thesis, and show that each theorist has slightly different assumptions in their theories. I then go on to show how decision-makers that hold different metaethical assumptions would end up applying those criminalisation theories differently, sometimes in ways that might go against the assumptions under which the theory came to be. I argue that it is therefore important and useful to be as clear as possible with metaethical assumptions made in theorising the deliberation process, particularly if one wishes those theories to actually be used by decision-makers at some point.

Chapter 5 focuses on a different kind of criminalisation theorising, which has been dubbed the 'political turn' in criminal law theory. The chapter summarises a few examples of this kind of theorising and applies the metaethical lens to them. The main takeaway of doing this is the fact that these theories depend on a metaethical distinction between normative domains of 'politics' and 'morality'. From this analysis informed by metaethics, I argue that if one wants to be this kind of theorist, it is important to address

the metaethical implications of making such a distinction, and that one must decide whether one's theory is a variation of legal moralism that focuses on political morality, or a more radical version that proposes politics as its own distinct normative domain.

CHAPTER 2 – METAETHICAL QUESTIONS FOR UNDERSTANDING CRIMINALISATION

Having now some clarity as to the importance of generally asking metaethical and meta-normative questions about how we understand criminalisation, I will now explain the specific questions that will be of interest for this thesis. I have called them the conceptual, the ontological and the epistemic questions, and I will explain them in detail below. In this chapter, I will begin by analysing the kinds of issues that form the conceptual question, particularly focusing on how wrongness gets changed by being 'moral' and how the definition of morality one assumes also defines wrongness. Next, I focus on the ontological question, showing some of the views that have been developed in metaethics on the possible ontological nature of moral properties. I will summarise the views of moral realism, moral anti-realism, moral naturalism and non-naturalism, as well as a brief discussion on metaethical constructivism. These tools will allow us to see how the ontological assumptions one makes about moral wrongness as a property of conduct can completely change how that kind of wrongness is understood. Finally, I focus on the epistemic question, where I discuss the possibility of there being such a thing as moral knowledge, depending on the metaethics one assumes. I then briefly summarise some views of moral epistemology, and discuss the importance of the epistemic question for practical deliberations, as it determines the epistemic method one uses to obtain moral knowledge.

2.1. The Conceptual Question

The conceptual question, as mentioned earlier, seeks to understand what conceptual use is being given to morality as a normative kind. Answering this

question involves determining what is distinctive or characteristic of the moral wrongfulness being used in the context of discussions about criminalisation. It also requires some basic definitions or delimitations of what counts as 'moral' in the context of this use, and what (if anything) is at play in that definition. That is, what kinds of assumptions and commitments are involved in the defining of the relevant sense of 'moral' in this context. In particular, I am interested in exploring how metaethical positions[29] can affect (if at all) how one defines what counts as morality, which therefore later affects how we use that understanding of morality in determining the moral wrongfulness relevant to criminalisation.

It is important to be clear that the conceptual question is not the monumental task of trying to define what morality actually is as a normative kind. Rather, my interest here is in which kinds of defining characteristics are being *used* or *attributed* to the specific concept of 'moral' wrongfulness. I wish to remain agnostic as to what actually defines morality 'correctly' and, instead, see how the usage of moral wrongfulness in the context of understanding criminalisation entails certain putative characteristics of morality as a normative kind. However, I will still make some suggestions, based on metaethical literature, of some success conditions for the conceptual use of 'moral wrongness', which I will argue lies in the way in which morality is attributed to have a particular kind of normativity.

[29] Some philosophers prefer to talk of *metamorality* for this kind of inquiry rather than metaethics, in order to distinguish the specific kind of positions that can be taken with regards to morality as separate from a more general ethics. For my purposes, however, though I am willing to grant that metamorality as a philosophical endeavour makes much sense, I will mostly prefer using the term metaethics, as the tools that are obtained from metaethics are still capable of being applied to morality - metamorality, in this sense, is a relevant sub-set of metaethics and metaethical positions can still be specifically applied to morality.

I will prefer this approach, rather than trying to provide an account of success conditions based on what morality is *about*. It is not a useful endeavour to try and define the distinctiveness of morality based on its target, in part due to its highly contested definitional nature.[30] More importantly, however, any actions or conducts that may be targets of morality can also be targets for other normative arenas. For example, if the target of morality were 'human flourishing', it seems plausible to state that considerations from prudence – which Fletcher characterises as "thought and talk about what's good/better/best or bad/worse/worst for, harms, benefits (etc.) someone"[31] – are also relevant for human flourishing. In addition, trying to find a specific property or characteristic of the targeted conducts — such as the common example of harmfulness of the conduct — misses the point that we may disagree on whether the particular characteristic is true or not (i.e. a moral disagreement) but that does not change the fact that we are talking about something moral.[32]

Asking these kinds of questions about the distinctness of morality involves a particular set of assumptions which are worth making explicit. I will assume that morality *is* the kind of thing that can be characteristic as a normative domain — that morality can, for example, have norms, properties, reasons and judgments that are distinct from those of other normative domains, like aesthetics, etiquette, rules of games, epistemology, prudence and so on. This also entails that there actually are distinct normative domains and that, though they may be inter-related, they are sufficiently definable as

[30] Michael Smith, *The Moral Problem* (Blackwell Publishers 1994) 39–41.
[31] Guy Fletcher, *Dear Prudence - The Nature & Normativity of Prudential Discourse* (Oxford University Press 2021) 1.
[32] See Walter Sinnott-Armstrong, 'The Disunity of Morality' in S Matthew Liao (ed), *Moral Brains: The Neuroscience of Morality* (Oxford University Press 2016) 335–338.

characteristic domains in themselves.[33] Assuming the contrary would entail either that there are no distinct normative domains as such, which seems ostensibly incorrect since it is clearly intelligible to see certain contexts, activities and frameworks as particular normative domains (like, for example, the game of chess); or it would entail that morality is not distinguishable as a normative kind. I will say something about this further below, but suffice for now to say that, at the very least, the uses and understandings that are given to morality suggest that there are certain aspects of moral normative concepts that are characteristic or distinctive in comparison to other potential normative systems.

It is also worth making explicit the assumptions I am making for why the distinction between senses of wrongfulness made earlier in chapter 1 is useful. There may sometimes be a temptation to understand 'wrongfulness' and 'moral wrongfulness' as synonymous. In other words, it makes no sense to distinguish between something being 'wrong' and something being 'morally wrong', as I am using them here, because that something is 'wrong' simply *means* that it is morally wrong. This would be because, the argument would go, something being 'wrong' without any other qualification of subscription to a particular normative domain would be a kind of wrongness that one *really* has to consider, and that is precisely what 'morally wrong' means.

The underlying assumption that needs to be made, as Gerritsen points out, for this idea to work is the commitment to ethical internalism, which "is roughly the view that morality necessarily or always involves a practical

[33] For some possible characteristics that can define a normative domain or system, see Derek Baker, 'The Varieties of Normativity' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2018).

dimension".[34] That is, if something is morally wrong, the view goes, then we necessarily have reasons to refrain from doing that action (reasons internalism) or we necessarily are motivated as agents to refrain from performing that action (motivation internalism). From there, one might take it a step further and, following Goldsworthy's distinction between substantial and conceptual internalism,[35] hold that it is conceptually true that morality has this kind of practical dimension. And, as Gerritsen rightly tells us, "only if conceptual internalism is true can we take the question about what is morally required and the question of what we have reason to do as one" and that this "is a specific thesis about the meaning of our moral concepts which can be rejected even by someone who does believe that we often or always have a strong reason to be moral."[36]

In view of this, conflating 'wrong' with 'morally wrong' implies that one is making this assumption from conceptual internalism, in the sense that what seems 'wrong' without qualification simply *is* what counts as morally wrong. I will discuss this conceptual issue of what counts as morality further below, but for now I wish to defend the prior idea that this conceptual conflation of wrongness with morality is not necessarily true. On the contrary, in order to avoid committing to a particular conception of what morality may be, it becomes necessary to make the distinction between the broad and narrow senses of wrongfulness I have presented here.

Returning to the conceptual question, the definition of what counts as moral wrongfulness, as a distinct kind of wrongfulness, will at least partially depend on how we answer the conceptual question of what we think counts

[34] Gerritsen (n 18) 13.
[35] Jeffrey Goldsworthy, 'Externalism, Internalism and Moral Scepticism' (1992) 70 Australasian Journal of Philosophy 40.
[36] Gerritsen (n 18) 15.

as moral. McPherson and Plunkett illustrate the point well: "when two philosophers each advance views about how to best explain what they each call 'moral' thought, talk, and reality, when are those rival explanations of the *same* part of thought, talk, and reality vs. when they are talking about different things, based on different meanings of the term 'moral'?"[37] The conceptual question helps in answering this question, and is particularly useful for understanding moral wrongfulness. As stated earlier, wrongfulness can be understood in a broad sense as any action that goes against a normative standard, so if we wish the narrower sense of wrongfulness (i.e. moral wrongfulness) to make any sense we need to be clear on what makes it distinct from the broader sense of wrongfulness and why.

A first important distinction – which may seem terminological but has important substantive consequences – is between 'ethics', 'morality', and '*mores*'. The differences between these concepts are important because they are the first step in defining which is the target of the wrongfulness we seek to find when criminalising a conduct. '*Mores*' is traditionally used to refer to the local, socially posited norms of behaviour that are relevant to a particular community of people – what is the wrong thing to do "according to them". 'Morality' in a stronger, normative sense,[38] refers to the norms that can transcend the locality of norms of *mores*. As Darwall puts it, morality "refers to an essentially normative order that stands in potential criticism of all positive *moralities*."[39] 'Ethics' and 'morality' are usually treated as terms that can be used interchangeably, but some philosophers – notably Bernard Williams – argue that there is a difference between the two. For Williams, morality is a distinct ethical kind, which is characteristic in its use of

[37] McPherson and Plunkett (n 2) 15.. Emphasis in the original.
[38] See Bernard Gert and Joshua Gert, 'The Definition of Morality', *Stanford Encyclopedia of Philosophy* (2020) 19–30.
[39] Stephen Darwall, 'Ethics and Morality' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2018) 554.

obligations and by the fact that these obligations are such purely in virtue of one being a certain kind of agent with certain capacities for agency, which are inescapable[40] and that characteristically elicit blaming as a reaction to not fulfilling a moral obligation.[41] Though one may disagree with Williams's particular characterisation, his underlying point is still useful to note – there is a distinction to be made between the broad sense in which we can have norms that help determine what to do, and the more narrow sense of a particular way of understanding things being right or wrong that purports to be of universal application. So, there are three potential levels of norms and normativity that we may be referring to when saying that something is 'morally' wrong: wrong according to *mores*, wrong according to morality or wrong according to ethics.

In most cases, it would seem that criminalisation literature is not intended to be read – when appealing to moral wrongfulness – as referring to *mores* or social morality. In fact, some authors explicitly say so: Duff says that his project is concerned, following Hart's terminology, with 'critical' rather than 'positivist' legal moralism.[42] The description of positivist legal moralism as that which looks at what is "thought to be wrong by members of this society",[43] fits with the description of *mores* presented above. Others are less overt about the issue, but one can infer that they are probably speaking of this 'critical' morality or, at least, not of *mores*. Simester and Von Hirsch, for example, speak of a truth-constraint on the state's moral voice used through the criminal law, and that the state should "get it right" when labelling

[40] Williams's use of 'inescapable' is slightly different to the one I will be using further below, in that his idea means that one cannot 'get out' of the normative domain of morality just because one's desires are incompatible with what morality demands.
[41] See Ch. 9 "Morality, the Peculiar Institution" in Bernard Williams, *Ethics and the Limits of Philosophy* (Routledge 2011).
[42] Duff, *The Realm of Criminal Law* (n 1) 53.
[43] ibid..

conduct as wrongful.[44] When they give examples of cases that can "straightforwardly be characterised as wrongful", they invoke situations like culpably harming someone, creating unwarranted risk of injury or flouting important communal obligations.[45] Though these situations could plausibly also be relevant to mores, the authors importantly do not speak of 'perceptions' or 'customs' as things that can establish wrongfulness, but instead say that criminalisation "makes an official valuation concerning the moral status of Φing".[46] This can be read as a purported reporting of 'actual moral facts' that will be true or false – as in the purportedly universal principles in the domain of morality – or as a reporting of 'perceived moral facts' of a particular community, where one might still mistakenly attribute a fact to a community's *mores*, but the mistake is on the descriptive fact of the community holding or not a particular view, rather than on whether the view is *actually* true. It seems that the authors are proposing their truth-constraint in the former sense, rather than in the sense of the descriptive facts about a community's mores. Just as a point of contrast, a historical example of someone who believed that *mores* was actually the target normative domain for the relevant wrongfulness for criminalisation was Devlin who, in his "The Enforcement of Morals", spoke of enforcing "the common morality" and that "the morality which the law enforces must be popular morality."[47]

One way to frame the difference between morality and *mores* that might be useful, in order to begin to argue which one might be relevant to criminalisation, is to look at which kind of fact is involved in each one and, in turn, which kind of fact we would need to identify and report back for the criminalisation process. In the case of *mores*, it seems that what one would

[44] Simester and Von Hirsch (n 1) 19.
[45] ibid 20.
[46] ibid 12.
[47] Patrick Devlin, *The Enforcement of Morals* (Oxford University Press 1965).

need to find is a series of descriptive facts about people's beliefs – if the concern is with what people in a particular community perceive or treat as wrongful, then the only fact of the matter one can identify within that domain is a social fact about what people are actually perceiving or treating as wrongful. The question of whether they are more generally 'right' or 'wrong' in *having* those perceptions or reactions is a question that belongs to a different normative order, and the answer seems to be a kind of moral answer – it is not on whether the report on *mores* wrongfulness is accurate, but on whether people who hold that belief are right in doing so. That being the case, if one believes that the 'moral' wrongfulness used within our criminalisation process is of this higher, 'critical' order, then that entails committing to looking for moral properties that are independent of – or at least different to – *mores* properties. To the contrary, if one believes that the 'moral' wrongfulness used in our criminalisation process is not of this sort, but rather is on the level of *mores* and local social customs regarding what is right and wrong, then one commits to looking for that kind of wrongness in conduct, rather than for something closer to a purportedly universal wrongness. The important thing to note, however, is that these differences come from conceptual differences in what the 'moral' in moral wrongfulness is supposed to mean, and that the choice of which concept to use is worth explaining.

There is also a need to explain whether one sees the 'moral' in moral wrongfulness as a *kind* of normative domain or as *the* normative domain *a-la*-ethics in Williams's terms shown above, as in morality being the only real, relevant normative domain for action-guiding decision-making. In this sense, we can understand the relevant wrongfulness as wrongfulness *according to morality* or according to moral norms – thereby entailing a delimitation of what goes into the moral domain and what norms are included in it (and,

logically, which are not) – or we can understand it as wrongfulness *according to ethics,* which seems closer to a kind of wrongness *simpliciter* or a genuinely all-things-considered wrongfulness.[48] Again, regardless of whether one agrees with Williams's distinction in particular, the underlying idea still remains relevant to this discussion: there are two senses in which something can be 'morally wrong', depending on the kind of domain one might think that wrongfulness finds itself in. One can mean wrongness as 'relative to morality', which makes it more discreet and limited to a particular domain, or one can mean moral wrongness as wrongness *simpliciter*, as the wrongness which one discovers after an all-things-considered deliberation.

So, in a sense, the distinction I made earlier is actually a reflection of how we delimit our conceptual understanding of morality. Therefore, it will be useful to pay a closer look to how one can go about doing so.

2.1.1. Delimiting and Defining the 'moral' in moral wrongness

Can we say that morality is distinctive as a normative kind? This question, in one form or another, has been present in moral philosophy since it began. However, it is important to be clear on what it is exactly that we are asking about – what is at stake in differentiating the 'moral' from the 'non-moral'? As stated earlier, since wrongfulness can be read as simply going against some normative standard of some kind, the narrower sense of moral wrongfulness must mean something distinct. One option might be, simply, that moral wrongfulness means going against the normative standards of morality, but if that is true, we end up in an unacceptable circularity: what does it mean to

[48] I say 'genuine' here in the sense that one can have a sense of 'all-things-considered' that means something like 'after all relevant factors in the deliberation process have been properly processed', rather than the seemingly thicker sense of 'all-things-considered' as a kind of normative domain in itself, which seems to carry with it a stronger sense of definitiveness in the resulting judgment of wrongfulness.

go against the normative standards of morality, in comparison to other possible normative standards? It seems trivially true to accept that moral wrongfulness means going against moral norms, but we have not learnt anything substantive about morality from that fact. And yet, it is a plausible position to hold to say that moral wrongness is wrongness relative to morality, as a contrasting normative kind with regards to other normative domains. So, something more needs to be said in order to properly understand what it actually means to do something wrong in a moral sense.

There are several potential answers one might give to the question on the distinctiveness of morality,[49] but for our purposes I am not pursuing one particular view of how morality *as a whole* is distinctive. I am more interested, rather, in determining how the concept of wrongfulness in particular is affected by the addition of a moral character to it. So, I will focus on identifying potential characteristics that might be added to (or taken away from) wrongfulness by the fact that it is considered a moral form of wrongfulness. First, I will present a few distinctions provided by metaethical analysis that will help in determining these potential characteristics.

The first distinction is between content-oriented and normativity-oriented characteristics of morality. By content-oriented, I mean characteristics that are associated to the norms that count as moral norms. By normativity-oriented characteristics, I mean those which are dependent on what specific kind of role the norms of morality are supposed to have within our normative lives. Here is an analogy that might make the distinction clearer.[50] If we were to look at morality as if it were a rulebook, content-oriented characteristics are those regarding the rules included in the book and what they are about;

[49] For some examples, see Ch. 1 in Smith (n 30)., especially 3-4.
[50] I borrow this analogy from Gerritsen (n 18).

normativity-oriented characteristics, in contrast, are about how we are supposed to use the rulebook itself, and the role it is supposed to play in our lives.

Further, within normativity-oriented characteristics one can make a distinction between two broad kinds of normativity that are observable: formal normativity and authoritative or robust normativity.[51] Formal normativity, as Baker puts it, "is the normativity displayed by any standard one can meet or fall short of."[52] The classic example of formal normativity are the rules of chess – there are right and wrong moves within the context of a chess game, one can state true or false facts about those rules and, if one does not follow the rules of chess while playing a game of chess, then one is doing something wrong relative to the normative standard of the rules of chess. In contrast, authoritative or robust normativity is somewhat more difficult to pin down conceptually. This is mainly because references to the 'authoritativeness' or 'robustness' of the normativity involved are expressed mostly through metaphorical or further normative language, such as talk of force, weight, importance, what really matters, and so on. One key aspect that distinguishes them, however, has to do with the way in which we understand the possibility of *ignoring* or setting aside the norms and reasons of each kind of normativity. For the purposes of this thesis, I will call this distinguishing characteristic between formal and authoritative or robust normativity *inescapability*.[53] Inescapability, in the sense I use it here, means that if there are authoritative or robustly normative norms or reasons involved in whatever

[51] See Tristram McPherson, 'Against Quietist Normative Realism' (2011) 154 Philosophical Studies 223, 232–233.. Both categories can be used to describe either norms (formal or authoritative norms) or reasons (formal or authoritative reasons).
[52] Baker (n 33) 568.
[53] The term has also been used for slightly different purposes by Foot in Philippa Foot, 'Morality as a System of Hypothetical Imperatives' (1972) 81 Philosophical Review 305., specifically to refer to the impossibility of escaping the requirements of morality by having or not having particular interests or desires.

we may be considering or facing, they should not be ignored unchallenged. This does not mean that authoritative norms and reasons always win out or override other norms and reasons,[54] but it does mean that they cannot be dismissed without alluding to some other legitimate norm or reason that explains not following the authoritative norm or reason.

This is precisely the kind of characteristic that one might want to ascribe to morality as a normative kind: morality, as a normative domain, has authoritative or robust normativity — in the terms used earlier, the norms and reasons of morality are inescapable. Again, there may well be reasons that can defeat moral considerations — they may be defeasible — but they cannot simply be ignored and not be taken into account at all. This is not the same as saying that moral considerations – or authoritative norms more generally – are always *relevant*. Relevance, the way I am using it here, refers to whether a particular norm or reason is linked to a particular type of activity or context for which the decision-making process is taking place. So, for example, if I am deciding whether to buy a particular ink for my fountain pen, a moral consideration like, say, 'do no harm' will not immediately and automatically be relevant to that decision. There *may* be situations when that consideration will *become* relevant – like, say, if I find out that the inks are made from turtle shells and sourcing them produces harm – but the consideration was not relevant from the start of my deliberation. Inescapability, in contrast, refers to the status of requiring a defeating reason for not being followed in a decision-making process – precisely one of those defeating reasons might be pointing to relevance to the decision-making context. In my example above, if I do find out that the ink is sourced as described, I cannot escape the fact that said

[54] For some examples of arguments that this is not necessarily the case for moral reasons, see Joshua Gert, 'Normative Strength and the Balance of Reasons' (2007) 1 Philosophical Review 533–562 and Douglas W Portmore, 'Are Moral Reasons Morally Overriding?' (2008) 11 Ethical Theory and Moral Practice 369–388.

consideration should go into my deliberation process, and I will have to provide myself with some defeating reason to not consider it. For example, I might really need that ink for an artistic project that will further my career, and so that gives me a defeating prudential reason to the moral 'do no harm' consideration. My point here is not to say that that is the correct or incorrect decision, but rather to show the deliberative mechanics of inescapability that I am suggesting for moral reasons: they are not always relevant, but when they are, they become inescapable and must be addressed in the deliberation. So, these two concepts are linked, but are not the same: relevance is more specific than inescapability.

Contrasting this characteristic to other normative domains might prove useful. Here is a specific example: suppose there is a fly standing on my chess board. All the pieces are placed in their starting positions and the fly is standing right in the middle of the board, but importantly, I am not in the middle of a game of chess – I am just practicing tactics on the board. Suppose further that I am truly irritated by the presence of the fly, and I wish to rid myself of its presence for good, as it is a particularly annoying fly. So, I am deciding what to do about it and one option I have is to grab a piece – let us say the king-side rook in its starting position – and squash the poor fly with it. Now, suppose for the sake of argument that we agreed that killing a fly is morally wrong in this context. Also, let us take as a given that, from its starting position, moving the rook to the centre of the board is an illegal move in chess. Does it make sense for me to think that I should not use the rook in this example because it would be an illegal chess move? Of course not, because I can easily escape from the normative context of a game of chess – the fact that it is an illegal move in chess is not a relevant reason, in this context, to the rightness or wrongness of using the rook to squash the fly and, therefore, I do not need to provide an explanation as to why I am not considering the fact

that it is a wrong move when deciding what to do. However, does it make sense for me to think that I should not use the rook in the example because it would be morally wrong to kill the fly? This does seem like a consideration I would have to take into account in my decision-making. It might, of course, not be a definitive reason against using the rook to squash the fly, but it is not a reason I can simply ignore or discard as irrelevant from the get-go. Importantly, if I did not take into account the moral property of the action of killing the fly being wrong – and, therefore, the moral reason that is created by that fact that I should not kill the fly with the rook – into my decision-making, it seems like I have not taken into account all the relevant facts about this state of affairs to make my decision. And it seems that, at least at some level, I have made an important mistake in not doing so – the kind of mistake for which I might be criticised, and for which I would probably be expected to explain why I have not taken that moral fact into account.

Now, suppose that we were actually in the context of a chess game and the same situation happened. In this case, I do have a reason to take into consideration the using of the rook, because I actually am in the context of a chess game and, therefore, if I use the rook I will be making an illegal move and might lose the game – perhaps I should use a different piece, say the king's pawn. However, could I do the same exercise I did before with the rules of chess with my moral consideration that it is morally wrong to kill the fly? It does not seem like this is possible - I cannot say that I can escape the authoritativeness of the moral reason just because I am now in the context of a chess game. In other words, the moral consideration still holds even if I am currently in a different normative context. Of course, my moral reason might still be defeated - I might have a very good reason to rid myself of the fly, even if I think it is morally wrong to kill it. But that is precisely the point: it needs to be *defeated*.

Suppose now, for the sake of argument, that we said there is nothing morally wrong with killing the fly, and that there is nothing morally right about it either – it is morally inert whether the fly lives or dies. But suppose that I still had an emotional response to killing the fly – I do not think it is morally wrong to kill it, but I do not enjoy the prospect of killing it as it will make me feel sad. These emotional considerations also seem escapable – it is perfectly plausible for me to think that this is not a time for emotional considerations or that I have a weightier reason (getting rid of the fly on my board which is making me lose concentration in this important chess game, for example) and, therefore, leave those reasons aside as irrelevant or defeated in this context. My emotional reason may either be defeated – I take it into consideration and another set of reasons win out – or I can simply escape the reason due to its irrelevance, and say to myself that emotional reasons have no role to play in my decision-making of squashing the fly during a chess game. By taking away the moral character of the reasons I may have for or against killing the fly, the reasons lose their inescapability. But if the reason does have a moral character — if, for example, there actually were a moral norm that told me I should not do things that make me sad — it demands attention, it requires consideration and it becomes authoritative regardless of the normative context in which a decision is being made.

If this proposed characteristic of morality seems plausible, then the implications for moral wrongfulness are now clearer: the fact that an action is 'morally' wrongful entails that the fact it is wrong is inescapable to anyone considering whether to perform the action or not, in view of the authoritative or robust normativity that morality brings with it. This is a conceptual claim of what it *means* for an action to be morally wrongful, and in order to get to this

conceptual claim we needed to use the tools provided by metaethics with regards to the nature of normativity.

One might, of course, disagree with this characterisation of moral wrongfulness as inescapable in the sense I have used it here, either because one thinks that morality need not be authoritatively normative or because inescapability – as I used it – does not seem like the correct way to portray authoritative normativity. The point here, however, is not to defend this particular characterisation of morality's distinctness as the correct view. The point, rather, is to show that this kind of characterisation affects how one understands the wrongfulness involved in criminalisation. And these understandings are important to be made explicit and addressed in our theoretical understandings of criminalisation. If by morally wrong we mean inescapably wrong, then the property of moral wrongness will look one way; if we mean something else – because we think 'morality' *means* something else – then that property will look a different way. This will also entail a difference in how we attribute that property to the candidate conduct for criminalisation, as well as the ways in which one can interpret the stating of a particular conduct as morally wrongful. So, ultimately, the conceptual questions regarding what counts as morality are importantly linked to our understandings of what criminalisation is doing, and of how we should go about criminalising conduct. Importantly, it can also allow us to better understand and evaluate those theories that use moral wrongfulness in some important way to guide or explain criminalisation, and to make sure that we are all talking about the same thing (or not).

It is useful, then, to focus on the normativity of morality as what makes it distinctive and trying to understand what that entails. To do this, metaethics is crucial, since it includes a rich debate around the existence of this kind of

authoritative or robust normativity.

2.1.2. Delimiting Authoritative Normativity

It is quite difficult to pin down exactly what authoritative normativity is supposed to be. It has gone by many names – robust, substantive, full-blooded normativity, among others – and the ways in which it is characterised is usually metaphorical or vague: what we *really* ought to do, what most forcefully ought to be done, what inescapably must be done, and so on.[55] What does seem clear, however, is that the common thread for all of these ideas lies in the importance that the contents of some norms should have to us, as agents and decision-makers. That is, authoritative normativity represents the highest level of importance to action-guidance that any kind of norm can have – hence why it is commonly referred to as what we *really* ought to do. If a norm has authoritative normativity, then it becomes important to, at the very least, take into account what that norm says when deciding what to do, think or believe.

But, as it turns out, the actual definition of what authoritative normativity is, perhaps not surprisingly, will be dependent on what metaethical commitments one has. In particular, the definition will depend on whether one thinks the concept of 'authoritative' is supposed to be agent-neutral or agent-relative.[56] That is, whether one believes authoritative normative reasons are dependent on the attitudes, roles, stances and commitments of the agent to whom the reasons apply (agent-relative reasons), or whether one believes that authoritative normative reasons do not depend on any of those

[55] For a detailed breakdown of possible understandings of authoritative normativity, see Ch. 2 in Gerritsen (n 18).
[56] For an in-depth explanation of the distinction, see Michael Ridge, 'Reasons for Action: Agent-Neutral vs. Agent-Relative', *Stanford Encyclopedia of Philosophy* (2017).

factors (agent-neutral reasons). In terms of what I am discussing in this thesis, is the fact that something is 'authoritatively' wrong – and any reasons that may derive from that fact – dependent on the commitments of agents or not? I say 'authoritatively' here rather than 'morally' only because, as we shall see below, it might make sense to think that morality is authoritatively normative, but I wish to leave open the possibility that other normative domains might also be authoritatively normative.

A useful way of visualising the question of authority in the context of normativity has to do with how a potential decision-maker is supposed to care about what a norm with authoritative normativity tells them to do or think. In other words, our definition of authoritative normativity should provide an answer to the question that a decision-maker could pose: 'why does this matter *to me*?' One potential answer is to say something along the lines of 'because this matters *to everybody*', which is close to the traditional idea of inescapability in the literature[57] – any agent, any potential decision-maker, is 'inside' the normative domain and its requirements, and they cannot get out of that position because *everybody* is within that position. As Enoch expresses it with regards to morality as being inescapable, "morality is not just a game that generates criteria of correctness – it is *the right* game, a game whose correctness conditions you can't escape by refusing to play it."[58] Authoritative normativity, however, seems to be broader than inescapability, because norms that have authoritative normativity can, as Joyce puts it, be "legitimately applied to a person irrespective of her ends."[59] In this sense,

[57] For some examples, see Foot (n 53); TM Scanlon, 'Contractualism and Utilitarianism', *Utilitarianism and Beyond* (Cambridge University Press 1982); Joseph Gilbert, 'Moral Inescapability' (1975) 10 Journal of Thought 8; RH Myers, 'The Inescapability of Moral Reasons' (1999) 59 Philosophy and Phenomenological Research 281.
[58] David Enoch, 'Non-Naturalistic Realism in Metaethics' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook in Metaethics* (Routledge 2019) 33.
[59] Richard Joyce, *The Evolution of Morality* (MIT 2006) 62.

following Joyce, authority can be understood as an imperative that "the subject would be irrational in ignoring it, or at least the subject has a reason of genuine deliberative weight to comply."[60] That is, it is less about not being able to get out of 'the game' in order for norms to not apply to an agent, but rather that not taking into account what those norms demand is a kind of substantial deliberative mistake – there is an important reason for or against a certain action that, if it is not considered in the deliberation process, makes the whole decision-making task incomplete or flawed. Here is where agent-neutral and agent-relative views come in – if authoritative normativity is agent-neutral, then *any* decision-making agent would be making this mistake, no matter what the particular context of that agent may be; if authoritative normativity is agent-relative, then authoritativeness of this kind will still be dependent on the particular roles, stances and attitudes of the decision-maker and, therefore, only leads to a mistaken decision-making process if it is incompatible with the roles, stances and attitudes of the decision-maker.[61]

As I warned in the introduction to this thesis, this is not the place to offer final answers to the questions posed by metaethics, so I will not try to defend a particular view about authoritative normativity. The point of this discussion has been instead to show the importance of asking the questions in the first place, and I hope that it has become clear at this point that answering the question of what counts as authoritative normativity can lead to very different views of what counts as morality. For even if we can agree that morality – and thereby moral wrongness – is authoritatively normative, what this latter

[60] ibid.
[61] For examples of each kind of view in moral theories, as well as the original first uses of 'agent-relative' and 'agent-neutral', see Derek Parfit, *Reasons and Persons* (1st edn, Oxford University Press 1986), 27 onwards.

characteristic entails varies greatly depending on what we think defines authoritative normativity.

Can we then say that authoritative normativity is the defining characteristic of morality? This seems partially correct, in the sense that the picture on what conceptually counts as 'moral' is still incomplete if we *only* appeal to authoritative normativity as the delimiting factor for morality – it would leave open the possibility that *all* authoritatively normative norms, and the reasons they create, are potentially 'moral'. This is certainly not the case, as it seems plausible that there are other normative domains – prudence, epistemology, and rationality, for example – that can also be authoritatively normative.[62] For example, the epistemic norms that tell us how we should form our beliefs, for example, can be seen as authoritative in the sense discussed here – they *really* tell us how to form our beliefs, they are important to our decision-making about forming beliefs, they might even be seen as agent-neutral norms that simply apply to *all* rational agents that form beliefs.[63] Another example: a norm of logic that forbids validly inferring the inverse of a conditional (if *p,* then *q*, therefore if *q,* then *p*) can plausibly be described as authoritatively normative in the sense described here – one *really* cannot validly infer the inverse of a conditional, and we therefore always have an authoritative reason not to infer the inverse of a conditional. Similarly, the epistemic norms that inform our inference-making also seem to be authoritative in this way – not following them or treating them as only applying to those who are partisan to knowledge seems like an incorrect way of understanding said norms. If someone makes a mistake in their process of inference, we do not think "well, they must not be playing the knowledge

[62] For a recent study of the authoritative normativity of rationality, see Benjamin Kiesewetter, *The Normativity of Rationality* (Oxford University Press 2017).
[63] For some useful discussions of the normativity of epistemology, see Conor McHugh, Jonathan Way and Daniel Whiting (eds), *Normativity: Epistemic and Pragmatic* (Oxford University Press 2018).

game of epistemology", and instead we might think something like that they are outright mistaken, regardless of whether they are 'playing the game' or not.

If the above is correct, then the conceptual definition of what counts as morality will need something additional to authoritative normativity to be complete. In other words, not only will a norm be of a moral kind when it is authoritatively normative, but also something about that authoritativeness needs to be distinct. One way to try to do this is to propose a kind of 'subject matter' for morality,[64] but the difficulty with this is that we can plausibly think of cases where there is an action that has properties relevant to more than one normative domain, even if they may seem 'characteristically' moral. For example, trying to argue that the subject of morality is community well-being would imply that other normative domains are not 'about' community well-being[65], or at least not in the same way as morality – if that is the case, the what is really doing the work for defining morality is not being about well-being, but instead is this 'not-in-the-same-way-ness' of being about well-being in contrast to other normative domains. Equally, trying to argue that morality's 'subject matter' is harms and the avoidance of them would mean that harms are not really the 'subject matter' of other normative domains, which clearly seems false – one can speak of epistemic harms,[66] aesthetic harms,[67] prudential harms[68] and so on in a way that is not incomprehensible. And, again, if one were to say that morality is characteristically about harms in a different way to other domains, then that difference is what is doing the

[64] See Joyce (n 59) 64–66.
[65] Or, more problematically, that those other normative domains can be reduced to morality at the end of the day.
[66] Simon Barker, Charlie Crerar and Trystan S Goetze, 'Harms and Wrongs in Epistemic Practice' (2018) 84 Royal Institute of Philosophy Supplement 1.
[67] Arnold Berleant, 'Negative Aesthetics and Everyday Life' (2011) 1.
[68] Tatjana Višak, 'Sacrifices of Self Are Prudential Harms: A Reply to Carbonell' (2015) 19 The Journal of Ethics 219.

work and not the fact of morality being about harms. Further, one can also speak of both harmless wrongdoing[69] and of permissible harms,[70] so attributing a single 'subject matter' to morality seems to give an incomplete picture of what morality actually does and how it is used.

To summarise, probably the best way to provide some kind of answer to the conceptual question is to argue for the authoritative normativity of morality – moral reasons are authoritative. However, this only gets us part of the way to a full understanding of morality, because we will still have to provide some kind of story for where that authority comes from, and how it is understandable. And this story is important to our understanding of criminalisation – particularly if we think moral wrongfulness plays a part in it – because there are so many potential answers to the conceptual question that we should be particularly careful in stating which conceptual understanding of 'morality' we are proposing. As Plunkett rightly puts it, "we would often do best *not* just to use the term 'morality' but to be *more specific* about which purported feature(s) of morality we care about. In short: Which of the various features that different philosophers associate with "morality" are the one we want to be focusing on in the context at hand?"[71] To do this, theorists would do well to delve into metaethical investigation and see which story of authoritative normativity is more plausible to them.

[69] See for example Volume 4 'Harmless Wrongdoing' in Joel Feinberg, *Moral Limits of the Criminal Law* (Oxford University Press 1984).
[70] Frances Kamm, *Intricate Ethics: Rights, Responsibilities, and Permissible Harm* (Oxford University Press 2007).
[71] David Plunkett, 'Robust Normativity, Morality and Legal Positivism' in David Plunkett, Scott J Shapiro and Kevin Toh (eds), *Dimensions of Normativity - New Essays on Metaethics and Jurisprudence* (Oxford University Press 2019) 113. Emphasis in the original.

2.2. Ontological Question

The ontological question is one of the central issues discussed in metaethics. It seeks to explain the metaphysical and ontological nature of moral facts, properties and reasons. This covers both whether those kinds of things exist in the first place – do moral facts or properties exist? – and if they do, what they are like. It includes, but it not limited to, concerns about whether these facts or properties are a 'part of reality' – which manifests as the divide between realists and anti-realists – and on whether, if those facts and properties exist, they are the kind of things that are found in nature – which manifests as the divide between naturalism and non-naturalism. When the debate is specifically about the moral realm, these divides are referred to as moral realism and moral anti-realism, and moral naturalism and non-naturalism. In the context of these divides, 'part of reality' is meant to be read as something whose existence is not dependent on the prior existence of a mind, which is why most of the time the realist/anti-realist divide is referred to as a debate around the mind-independence of moral facts and properties.[72] As Joyce puts it, "traditionally, to hold a realist position with respect to X is to hold that X exists in a mind-independent manner […] On this view, moral anti-realism is the denial of the thesis that moral properties […] exist mind-independently."[73] In turn, the divide between naturalists and non-naturalists is with regards to whether moral facts and properties are, or can be reduced to, the same kind of facts and properties that are obtained

[72] For a criticism of this view, see Connie S Rosati, 'Mind-Dependence and Moral Realism' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2017).
[73] Richard Joyce, 'Moral Anti-Realism', *Stanford Encyclopedia of Philosophy* (2016).

through the scientific method – facts of the 'natural sciences' – or whether they are of a different, non-reducible and non-natural kind.[74]

Applied to moral wrongfulness specifically, the ontological question will ask whether moral wrongfulness is a mind-independent property of a conduct or not, and whether the property of moral wrongfulness is something that can be reduced to a natural fact or not. As one might imagine, there are several potential answers to both kinds of questions, the combinations of which generate different metaethical views – and potential commitments – on the nature of moral wrongfulness, which I will explore below. Thus, by asking the ontological question we can get a better picture as to what moral wrongfulness looks like *as applied to* criminalisation: when we say that a criminalised conduct conveys the conduct as morally wrong, are we referring to a mind-independent fact that is part of the conduct talked about, or are we referring to a mind-dependent fact that the speaker is reporting with regards to their own or somebody else's mind? When moral wrongfulness is proposed as a principle to include in the deliberation whether to criminalise a conduct, do we mean that the wrongfulness involved is a natural fact that can be observed through the scientific method, or do we mean a non-natural fact that needs to be accessed by some other means? Or, perhaps, we want to say that actually talking of moral wrongfulness is not referring to anything in the first place, and instead we should be talking about something else? These are all examples of issues that, if one wants to deal with them, one needs to answer the ontological question first.

[74] For in-depth overviews of both positions, see Matthew Lutz and James Lenman, 'Moral Naturalism', *The Stanford Encyclopedia of Philosophy* (Spring 2021, Metaphysics Research Lab, Stanford University 2021) <https://plato.stanford.edu/archives/spr2021/entries/naturalism-moral/> accessed 28 June 2023; Michael Ridge, 'Moral Non-Naturalism', *The Stanford Encyclopedia of Philosophy* (Fall 2019, Metaphysics Research Lab, Stanford University 2019) <https://plato.stanford.edu/archives/fall2019/entries/moral-non-naturalism/> accessed 28 June 2023.

Note that when asking ourselves about the ontological nature of moral wrongfulness, there are two potential targets to the question: either we are asking about the nature of the *property* that makes the conduct morally wrongful or we are asking about the nature of the *normativity* of that wrongfulness. That is, we can either look at the property itself under which a conduct is categorised as morally wrongful or at the kind of normativity that results from the moral wrongfulness in question. In other words, what are the ontological characteristics of the alleged moral property of an action that make it morally wrongful – whereby we will also be asking about the relationship between that property and moral norms – or, in contrast, what are the ontological characteristics of the reasons that the presence of a moral property elicits, which lead to the normativity of moral wrongfulness (i.e. how important to decision-making are the reasons moral wrongfulness generates)?

It is also not an ontological question about moral norms as entities. Regarding the ontological nature of a moral norm *qua* norm, it seems relatively uncontroversial to accept that moral norms – like any other potential norm – are manifested as a kind of rule or principle through theorising and that, therefore, their ontological nature might be purely conceptual or something along those lines – they are the result of theorising about what a moral norm would contain *based on* the existing moral properties being discussed. In other words they are, as Koller puts it, "*ideal* entities that have a propositional content expressing some normative aims or claims",[75] for which "it must, in principle, be possible to express them in the form of linguistic utterances, which can be represented through *prescriptive*

[75] Peter Koller, 'On the Nature of Norms: On the Nature of Norms' (2014) 27 Ratio Juris 155, 157. Emphasis in original.

or *norm-expressive sentences.*"[76] The ontological question I pose here does not focus on these ideal entities manifested as linguistic expressions. So, for example, the interest is not in the ontological characteristics of a moral norm like 'thou shalt not kill', but instead is in the ontological nature of the alleged properties of the act of killing that *make it the case* that 'thou shalt not kill' is a (putative) moral norm.

Answering the ontological question is what allows us to see what kind of metaethics one is committed to, because there are a large number of potential answers one might give depending on what one believes is the case regarding moral properties and reasons. This is because, traditionally, one goes about determining which metaethical school of thought one follows depending on how one answers a series of ontological questions about the properties that moral language refers to, and there is much space for diverging views at each stage of these series of questions. In fact, sometimes the potential answers one might give to these sets of binary questions is shown as a flow-chart,[77] though there have been alternative hybrid approaches between contrasting ontological positions explored in recent years.[78] Nevertheless, there does seem to be a series of questions about the ontology of moral properties that can place us in metaethical "camps", which is useful to be aware of in view of what I have said previously – there is a strong possibility that if we have profound disagreements as to the ontological nature of morality, that we will end up talking past each other when discussing things like moral wrongfulness.

[76] ibid. Emphasis in original.
[77] See, for example, Miller (n 20) 8.
[78] For an overview of the contrast between both these approaches, see Mark Schroeder, 'Hybrid Expressivism: Virtues and Vices' (2009) 119 Ethics 257; Guy Fletcher and Michael Ridge (eds), *Having It Both Ways: Hybrid Theories and Modern Metaethics* (Oxford University Press 2014).

For the purposes of this thesis, I will explain some broad ontological issues that will help us understand the main differences in views about moral properties, though I will not try to defend any particular view. Rather than serving as an exhaustive exploration of the intricacies of each metaethical position, my interest here is to show some of the potential positions one might take when talking about moral wrongfulness, and to understand the basics of each of those positions. I will begin by explaining the differences between moral realism and anti-realism, and then explain the differences between moral naturalism and non-naturalism. The purpose of this exposition is to then have enough tools to see how we might propose potential answers to the issues associated to the ontological question, and in particular to see how a theorist of criminalisation might view the ontological nature of moral wrongfulness.

2.2.1. Moral Realism

One possible answer, then, comes from the moral realist camp. It is complicated to pin down what being a moral realist entails, because what counts as 'real' for realism is contested. As Railton tells us, "General characterizations of realism are frustrated by the fact that some of the dominant images associated with realism about one subject matter do not apply readily, or at all, to realism about others."[79] One way in which the realism relevant for morality has been traditionally depicted has to do with the relationship between the existence of what one is talking about – in this case, moral properties – and its dependence on the further existence of a mind and its mental states. Broadly speaking, though not every kind of moral

[79] Peter Railton, 'Moral Realism: Prospects and Problems' in Walter Sinnot-Armstrong and Mark Timmons (eds), *Moral Knowledge? New Readings in Moral Epistemology* (Oxford University Press 1996) 51.

realist needs to agree with this,[80] as Rosati tells us, "the more common view, however, seems to be that moral realism requires mind-independence".[81] Therefore, specifically for the purposes that interest this thesis, a moral realist will usually make the claim that moral wrongfulness is a mind-independent property of actions and that, therefore, when we talk and think about moral properties and reasons, we are referring to things that are independent of our minds. The moral realist position usually includes the commitment that moral properties are facts that are 'out there' in reality, and their existence does not depend on a mind making it so. As Brink puts it, the realist wants to say that "the facts about the world are *not constituted* by the mental"[82] which, without wanting to get into a lengthy debate about what constitutive means, can roughly be understood as the idea that moral facts do not lose any of their ontological nature if there are no mental states present. In other words, if we were to take minds out of the equation of what moral facts are, those moral facts remain the same because they are mind-independent and not constituted by the mental.

Another way of putting the moral realist position has to do with a feature of moral *language*, in the sense that realists will want to say that moral statements are *truth-apt*. This means that moral claims – in our case, claims about moral wrongfulness – are the kind of claims that *can* be either true or false. And the realist will believe this is so *because* they refer to actual properties that exist in the world mind-independently. As Korsgaard puts it when referring to what she calls 'substantive' moral realism, it "is the view that there are answers to moral questions *because* there are moral facts or truths,

[80] For one example of a realist view that treats moral facts as 'standpoint-dependent', see David Copp, *Morality, Normativity & Society* (Oxford University Press 1995). I shall discuss these kinds of views further below, when discussing constructivism.
[81] Rosati (n 72) 356.
[82] David O Brink, *Moral Realism and the Foundations of Ethics* (Cambridge University Press 1989) 15. Emphasis in the original.

which those questions ask *about*."[83] The semantic point made here by the realist is the basis for the more traditional distinction in metaethics between cognitivism – the view that moral statements are truth-apt – and non-cognitivism – the view that moral statements are not truth-apt and are, instead, statements of conative attitudes that are neither true nor false.

However, note that the semantic point on moral language is still linked to what one believes is ontologically the case about moral facts, and we can inform our theories of moral language with that ontological nature. They are views that work together in the case of the moral realist and, as Chrisman puts it, "as soon as one thinks that ethical sentences represent facts obtaining objectively *and some of these sentences do so correctly,* then one can be classified as a metaethical realist of some sort."[84] It is also the case that there are now hybrid theories that propose that moral language might be *both* asserting truth-apt statements of fact *and* expressing conative attitudes of the speaker[85], and therefore one can eventually be either a moral realist or anti-realist ontologically speaking about moral facts, but not necessarily be exclusively a cognitivist or a non-cognitivist about moral language.

The way in which moral language is used is useful in figuring out what realist position is being proposed, because an important point that any moral realist will want to make is that moral language *refers* appropriately to things in the world, which thereby entails that there are those things in the world in the first place. This distinction also will become useful to understand certain forms of anti-realism that I shall explain below, that will agree with the realist that moral language *does* refer to things in the world, but it always *fails* to do

[83] Christine Korsgaard, *The Sources of Normativity* (Cambridge University Press 1996) 35.
[84] Matthew Chrisman, *What Is This Thing Called Metaethics?* (Taylor & Francis Group 2016) 33 <http://ebookcentral.proquest.com/lib/ed/detail.action?docID=4747553> accessed 19 November 2022.
[85] For a survey of these kinds of theories, see Fletcher and Ridge (n 78).

so successfully. In addition, the stronger versions of moral realism will also want to say that the appropriateness of this reference is independent of our opinions, stances, commitments, attitudes, desires and, more generally, our mental states – the fact that we may have particular instances of these mental states does not change whether the property we are referring to through our moral language is true or false. This 'stronger' version of moral realism fits with Eklund's proposed *ardent* realist, who will want to say when faced with someone (that Eklund presents as "Bad Guy") who "does bad things and is motivated by bad desires – is somehow objectively out of sync with how to conduct one's life: in no way are our differences with Bad Guy *merely* a disagreement in taste, or a matter of having different desires."[86] Additionally, Boyd proposes three basic theses that a moral realist believes in:

"1. Moral statements are the sorts of statements which are (or which express propositions which are) true or false (or approximately true, largely false, etc.);

2. The truth or falsity (approximate truth…) of moral statements is largely independent of our moral opinions, theories, etc.;

3. Ordinary canons of moral reasoning – together with ordinary canons of scientific and everyday factual reasoning – constitute, under many circumstances at least, a reliable method for obtaining and improving (approximate) moral knowledge."[87]

I have explained 1 and 2, and most if not all moral realists will agree with them, but Boyd's point 3 is contested and will require more explanation further below, as it is the distinction between moral naturalism and non-

[86] Matti Eklund, *Choosing Normative Concepts* (Oxford University Press 2017) 1.
[87] Richard Boyd, 'How To Be a Moral Realist' in Stephen Darwall, Allan Gibbard and Peter Railton (eds), *Moral Discourse & Practice - Some Philosophical Approaches* (Oxford University Press 1997) 105.

naturalism. For now, however, it is enough to point out that the moral realist will endorse points 1 and 2 about moral language precisely because they believe, in an ontological sense, that there exist moral properties and reasons to which moral language can accurately refer, and that the expressions of which can be true or false.

Therefore, if we were moral realists *and* believed that criminalisation states moral wrongfulness or uses it in deliberations about criminalising, then that will usually mean accepting that said moral wrongfulness is a mind-independent property of the conduct being criminalised. I will explore the implications of this in chapter 3 in more detail, but it is useful to keep these ideas clear from the start.

2.2.2. Moral Anti-Realism

A different answer to the ontological question comes from the anti-realist camp of metaethics. This view is the negation of moral realism, and therefore proposes that moral facts, properties and reasons are not mind-independent or stance-independent. This view can be defended in two ways: either by denying the existence of moral properties in the first place, or by accepting that they do exist but that they are, in some substantive way, mind-dependent.[88] The former strategy is embodied most prominently in moral error theory: the position which states that in view of there not being any moral facts or properties that actually exist, all instances of moral talk[89] are systematically false. The latter can manifest under different names, but is consistent with what Joyce calls 'non-objectivism'[90] and with some forms of

[88] See Joyce (n 73).
[89] Specifically, all instances of substantive or positive moral talk, like 'it is morally wrong to steal' or 'it is morally permissible to steal' - what Joyce calls 'atomic moral statements' which presuppose the instantiation of a moral property. See ibid 19.
[90] ibid 27–40.

constructivism, where the common view of both positions is that the ontological nature of moral properties is best understood as mind-dependent. For the non-objectivist, this means that moral properties are, in some important sense, *constituted* by some kind of mental phenomena,[91] whereas for the (anti-realist) constructivist it means that moral properties *come to exist* as a result of being entailed from a particular practical point of view.[92]

There is also, as referred to briefly before, the possibility of giving a non-cognitivist answer to the ontological question: if one takes the position that moral language is not truth-apt in the first place, it makes sense to also believe that this is because there are no moral properties to refer to through that language. Or, put differently, it seems very difficult to hold the position that moral language is not truth-apt, and yet there are mind-independent moral facts that moral language could potentially refer to, but simply does not. A historically relevant version of non-cognitivism was A.J. Ayer's emotivism,[93] which has sometimes been dubbed the "Boo/Hurrah!" theory, where moral expressions such as "Murder is wrong" are best explained as expressions of the speaker's disapproval of murder, in a similar way as saying "Boo to murder!". More modern versions of non-cognitivism, such as Allan Gibbard's norm-expressivism,[94] state that the non-cognitive, non-truth-apt expression that is happening in moral talk is the expression of an agent's acceptance of certain norms.[95]

[91] Joyce (n 73).
[92] Here I follow Street's definition of constructivism found in Sharon Street, 'What Is Constructivism in Ethics and Metaethics?' (2010) 5 Philosophy Compass 363.
[93] Alfred Jules Ayer, *Language, Truth and Logic* (Dover Publications 2012 [1946]).
[94] Allan Gibbard, *Wise Choices, Apt Feelings* (Oxford University Press 1990).
[95] For a detailed analysis, see Miller (n 20) ch 5.

Another important line of thought to mention is moral error theory. This position, originally developed by J.L. Mackie,[96] proposes that as a matter of fact, there is no such thing as moral properties or reasons (or, at least, not in the way realists argue), and that therefore moral statements, though they are truth-apt (i.e. following cognitivism about moral language) are uniformly false, because there exists nothing in the world that can instantiate what moral language attempts to refer to. As Olson puts it, "moral error theory [is] the view that moral thought and discourse involve systematically false beliefs and that, as a consequence, all moral judgments, or some significant subset thereof, are false."[97] Mackie bases his view on two main arguments: the argument from *relativity* and the argument from *queerness*. The first is the observation that the large amount of disagreement that exists between different moral codes is better explained by the "hypothesis that they reflect ways of life than by the hypothesis that they express perceptions, most of them seriously inadequate and badly distorted, of objective values."[98] The second, more important argument from queerness, is the conjunction of two ideas that Mackie puts forward thusly:

> "If there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe. Correspondingly, if we were aware of them, it would have to be by some special faculty of moral perception or intuition, utterly different from our ordinary ways of knowing everything else."[99]

From these characterisations of moral anti-realism, we can also get a clearer picture of a particular kind of moral realism: if one takes the position

[96] JL Mackie, *Ethics - Inventing Right and Wrong* (Penguin Publishing 1977).
[97] Jonas Olson, 'Error Theory in Metaethics' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2019) 58; Jonas Olson, *Moral Error Theory - History, Critique, Defence* (Oxford University Press 2014).
[98] Mackie (n 96) 37.
[99] ibid 38.

that moral language is truth-apt (cognitivism), *and* that there are moral facts to which moral language can refer and which make (at least some) moral statements true (denying moral error theory), *and* that those moral facts are mind-independent (denying non-objectivism and certain forms of constructivism), then one takes the position of what has been called *robust* moral realism.[100] That is to say, a robust moral realist will hold the position that in view of there existing mind-independent moral properties, it makes perfect sense to see moral statements as utterances that are truth-apt and that, if they correctly refer to these mind-independent moral properties, they are true statements. The key point to note for the ontological question I am posing, however, is that both the realist – robust or otherwise – and the anti-realist base a lot of their position on what they believe is the correct description of the metaphysical character of moral properties.

So, if we apply an anti-realist position to criminalisation, we can see that the entailments are quite different to those of moral realism stated earlier. To use examples regarding moral talk, if one follows some form of non-cognitivism – and accepts the idea that criminalisation portrays the criminalised conduct as morally wrongful – then what is being said when a conduct is criminalised is quite different: it is a report of a conative attitude that a speaker has about the conduct being criminalised, rather than a description or report of a particular state of affairs with regards to the conduct in question. In contrast, if one is a moral error theorist, what is stated through criminalisation is a purported description of states of affairs regarding the conduct being criminalised – it just uniformly states falsities about the conduct as there are no actual properties for the statement to refer to. Finally, if one is an anti-realist constructivist, what is being stated when using moral wrongfulness is something like that from our practical standpoint as decision-

[100] Joyce (n 73) 28.

making agents, and in view of certain values one has committed to, the conduct in question does not stand to scrutiny as acceptable in view of those constructed commitments we have made as judgers, instead of a mind-independent matter of fact about the conduct itself. I will explain these views in more detail further below when discussing constructivism more specifically.

As we can see, these potential answers differ greatly to those that a moral realist would agree to. In view of these differences, it perhaps now becomes even clearer that it is better to be clear about one's metaethical commitments when dealing with moral wrongfulness, because we might be thinking of completely different things. Of course, one might reply to this that maybe this is unnecessary, because most people might simply agree implicitly in their metaethics. This seems to be a common belief with regards to moral realism, in the sense that it is assumed to be the default position about morality and, therefore, it is only up to those who are sceptical of moral realism to provide evidence to the contrary.[101] Though there is little empirical evidence to suggest this,[102] even if this potential intuition were true, there are further differences even *among* moral realists, reflected in the divide between moral naturalists and non-naturalist which, as we shall now see, give us further reasons as to why it will be useful to be explicit with our metaethical views when discussing moral wrongfulness.

2.2.3. Moral Naturalism and Non-Naturalism

[101] For why it might be seen as the 'default position' to take, see Enoch (n 58) 30–33. As to why this might not be the case, see the supplementary entry "Moral Anti-realism vs Realism: Intuitions" in Joyce (n 73).

[102] If anything, it seems that people have inconsistent 'folk' metaethics, which is reflected in the results of contemporary empirical research on the matter, such as in Ross Colebrook, 'The Irrationality of Folk Metaethics' (2021) 34 Philosophical Psychology 684.; and in Thomas Pölzler and Jennifer Cole Wright, 'Anti-Realist Pluralism: A New Approach to Folk Metaethics' (2020) 11 Review of Philosophy and Psychology 53.

Additionally to the metaethical divides already explored, there is a further distinction to be made between those who accept that moral properties are factual, in the sense that there are such a thing as moral facts of some kind. This discussion will mostly be relevant for this thesis when we come to discussing moral epistemology, but that has its basis on the ontology of moral facts, so I present the distinction here. It is usually a divide applicable to moral realists, but the divide can also apply to some forms of constructivism. That is, moral facts might exist in a mind-dependent way, and those mind-dependent facts might be of a natural kind[103] or of a non-natural kind.[104]

Non-naturalism is the view that moral properties are not the kind of facts that are suitably obtained through something like scientific inquiry. As FitzPatrick states, "ethical properties (such as moral rightness or goodness) and facts (such as the fact that an act is wrong, or that a certain consideration is a reason for acting) are neither among the properties and facts that are the proper subject of scientific inquiry, nor constructible from those that are. They are instead *sui generis*."[105] Perhaps one of the most prominent contemporary non-naturalist moral realists, David Enoch, explains non-naturalism in the context of moral realism as follows: "the non-naturalist realist thinks of moral facts, properties, and objects as ontologically exciting at least in the sense that they are unlike facts about the average weight of the male, middle-aged analytic philosopher: They are not reducible to or entirely grounded in other, non-moral facts, properties and objects."[106] In other words, moral properties

[103] For an example of this, see Sharon Street, 'Constructivism about Reasons' in Russ Shafer-Landau (ed), *Oxford Studies in Metaethics*, vol 3 (Oxford University Press 2008).
[104] A good example of this latter strand of non-naturalism is found in Korsgaard (n 83)., where the constructivist standpoint that gives correctness to moral facts is still a *sui generis,* non-natural source.
[105] William J Fitzpatrick, 'Ethical Non-Naturalism and Normative Properties' in Michael Brady (ed), *New Waves in Metaethics* (Palgrave Macmillan 2011) 7.
[106] Enoch (n 58) 30.

are distinctly unique in comparison to other kinds of facts, and are not linked to any kind of facts that are found through empirical means.

In contrast, naturalism states that moral facts are the kind of facts that can be obtained by – or in some way reduced to – scientific inquiry, as they are part of the natural world. As one might imagine, defining what counts as 'natural' is contested – as Lutz and Lenman put it, "What does the *naturalness* of a fact or property consist in, exactly? That's a pretty basic question, but philosophers are often left embarrassingly tongue-tied when called on to answer it."[107] But we can make some basic approximations to an understanding. First of all, naturalism for moral properties proposes that they are knowable – in the sense of obtaining true beliefs about them – through some form of empirical process. This does not mean that moral *concepts* might not have some form of analytic truth, but rather that the ultimate truth-value of the moral properties themselves is only verifiable through some kind of empirical access. As Copp explains it, "the naturalist should say that any knowledge we have of synthetic moral truths must be empirical. And, in line with this, she might propose that moral properties are empirical in that any knowledge we have of synthetic propositions about their instantiation must be empirical."[108] This is precisely why the naturalist/non-naturalist divide will be relevant, in the context of this thesis, for the epistemic question: it will help understand what *type* of fact one is searching for when identifying moral properties such as moral wrongness.

One central difficulty that non-naturalist theories need to face, and that I will discuss regarding the epistemic question, are what Gerritsen calls the 'epistemological puzzle' they come with. As she puts it, "if normative

[107] Lutz and Lenman (n 74).
[108] David Copp, 'Why Naturalism?' (2003) 6 Ethical Theory and Moral Practice 179, 181.

properties are non-natural properties, which are often defined as being causally inert, how can we know about them? How could we have access to them?"[109] Answering this will require discussing epistemic methods for moral knowledge, but I still mention this issue here regarding the naturalist and non-naturalist stances that one can take.

2.2.4. Constructivism as a Metaethical View

One final metaethical view I will discuss is constructivism for morality. An important question that inspires constructivism as a metaethical view is the following: if it makes sense to think about there being moral facts, are those facts 'out there' for us to discover or have we 'put them there' from within, as a kind of construction? This question is linked to the mind-dependence issue of the ontological question: if we accept that there are moral facts, that they are a part of reality in some important way and that we can gain access to them, then are those facts mind-independent or not and, depending on the answer to this last point, is our knowledge about those facts trying to fit the world independently of the mind or not. If one answers this question by saying that there *are* moral facts, that moral statements *are* truth-apt, but that these are *not* mind-independent, then one is thinking in terms of constructivism.[110]

Constructivism is, broadly speaking, the view that moral properties are facts whose source is not mind-independent. The view of constructivism has been developed in political philosophy and normative theorising, famously

[109] Gerritsen (n 18) 118.
[110] This kind of view might also be consistent with a form of 'response-dependent' accounts of morality, which I explore briefly in chapter 3.

in the work of Rawls[111] and Scanlon,[112] but for it to stand as a view in metaethics, as Chrisman puts it, "one needs to claim that *all* ethical facts are 'constructions' rather than apt for discovery."[113] From there, as Barry explains, two claims emerge for constructivism about metaethics: "(1) substantive reasons and values are a function of procedurally correct reasoning from an agent's commitments, and (2) the procedures of construction are principles constitutive of agency, and applicable because of this."[114] Thus, for the constructivist, the source of moral properties is the result of a deliberation done by agents, which more recent versions of metaethical constructivism have dubbed the "practical standpoint".[115] This latter concept is characterised as a formal view of the standpoint of valuing in itself, which allows for an explanation of what it means to value anything in the first place. The important point to notice is that the existence of values, under this view, depends entirely on the existence of agents taking the practical standpoint. As Chrisman tells us, "According to the metaethical constructivist, if there were no beings who could take some things to be valuable, then nothing would be valuable. For there to be reasons to act, there must be agents taking up the practical standpoint, trying to decide how to act."[116]

There are, in turn, different ways in which metaethical constructivism can be developed. Street dubs these the 'Kantian' and the 'Humean' versions of

---

[111] John Rawls, *A Theory of Justice* (Harvard University Press 1971).
[112] Scanlon (n 57); TM Scanlon, *What We Owe to Each Other* (Harvard University Press 1998) <http://www.jstor.org/stable/j.ctv134vmrn> accessed 29 June 2023.
[113] Chrisman (n 84) 152.
[114] Melissa Barry, 'Constructivism', *The Routledge Handbook of Metaethics* (Routledge 2018) 392.
[115] See Korsgaard (n 83); Christine M Korsgaard, *The Constitution of Agency: Essays on Practical Reason and Moral Psychology* (Oxford University Press 2008) <https://doi.org/10.1093/acprof:oso/9780199552733.003.0008> accessed 29 June 2023; Street (n 103); Street (n 92). For critiques, see Carla Bagnoli, *Constructivism in Ethics* (Cambridge University Press 2013)
[116] Chrisman (n 84) 152.

constructivism,[117] where the former accepts the idea, regarding moral conclusions, that "we may start with a purely formal understanding of the attitude of valuing, and demonstrate that recognizably moral values are entailed from within the standpoint of any valuer as such."[118] In contrast, Humean constructivists "deny that substantive moral conclusions are entailed from within the standpoint of normative judgment as such. Instead, these views claim, the substantive content of a given agent's reasons is a function of his or her particular, contingently given, evaluative starting points."[119] In other words, the Kantian constructivist will find values embedded into the practical standpoint, whereas the Humean constructivist will not, and instead those values "must ultimately be supplied by the particular set of values with which one finds oneself alive as an agent – such that had one come alive with an entirely different set of evaluative attitudes, or were mere causes to bring about a radical shift in those attitudes, one's reasons would have been, or would become, entirely different."[120]

This brief summarising of ontological views on metaethics has not been intended to be exhaustive, nor to be a defence for any of these views. It is rather an attempt to give a taste as to the varied options that can be taken regarding the ontological nature of moral wrongness, and in showing this variety to make the case for being aware of them, and refer to them more explicitly, in our theorising of criminalisation.

---

[117] Street (n 92) 369–370.
[118] ibid 369.
[119] ibid 370. For criticisms of Korsgaard's view, see Nadeem JZ Hussain and Nishi Shah, 'Misunderstanding Metaethics: Korsgaard's Rejection of Realism' in Russ Shaffer-Landau (ed), *Oxford Studies in Metaethics*, vol 1 (Oxford University Press 2006).
[120] Street (n 92) 370.

2.3. Epistemic Question

Finally, the epistemic question concerns access to, and the existence of, moral knowledge. It is specifically interested in understanding the ways in which we gain access to knowledge about moral properties and facts, as well as in how we can put ourselves in the right epistemic positions in order to do so. So, answering the epistemic question requires that we first have an answer to the ontological question – we cannot determine what epistemic processes are needed to access knowledge if we do not know where to look for it in the first place. Once we have settled – or at least committed to – an answer to the ontological question, we can then move on to the epistemic question and try to determine how knowledge about those putative moral facts and properties works.

One potential answer, as we have seen when looking at the ontological question, is that there is no possibility of moral knowledge – since there are no moral facts or properties to gain access to – and so the very simple answer to the epistemic question might just be that you cannot have access to moral knowledge. However, assuming that this ontological description of moral facts were true, this does not necessarily mean that we could not have some kind of epistemic access to *other* kinds of normative facts or properties, of which one might say that there is such a thing as knowledge. So, from this perspective, the epistemic question might say that we cannot access moral knowledge, and we might take that answer as a reason to search elsewhere in the normative spectrum, such as prudence, social norms, or any other normative domain that one argues to be relevant, where we will end up asking the epistemic question again about a different normative system.

The epistemic question is particularly salient for criminalisation theories that propose using moral wrongfulness as a guiding principle of some kind for criminalisation. This is because I take it that criminalisation theories are intended to be *used* by potential criminalisation decision-makers – they are not only an explanatory academic exercise and are instead supposed to provide useful mental tools for when one is faced with a decision on whether to criminalise a conduct or not.[121] If this is true, one potential question a decision-maker might have goes as follows: if I am supposed to use moral wrongfulness as, say, a constraint on my decision to criminalise, how do I go about determining when a conduct has the property of being morally wrongful? In other words, *how do I find this out?* The answer to the epistemic question is precisely how we can give that methodological guidance for the decision-maker.

Note, however, that this is not the same as asking for a normative principle under which the moral property of wrongfulness obtains – this is a question of normative ethics rather than of epistemic access to moral knowledge. That is, the answer to 'how do I find this out' that is relevant to the epistemic question is not, say, that you must calculate the amount of utility or disutility that the conduct creates, or that you need to identify whether a moral duty has been breached. Instead, the relevant kind of answer is that you will need to *perceive* moral wrongfulness as a property of the conduct, or you will need to have a *dialectic exchange* with other people about what their responses are to the conduct, or you will need to look at your *institutional commitments* in order to determine, through some kind of rational reconstruction, what kinds of conduct fall within the category of wrongful relevant to your

[121] For example, Moore, *Placing Blame - A General Theory of the Criminal Law* (n 1) ch 16. is entitled "A Non-Exclusionary Theory of Legislative Aim: Taking Aim at Moral Wrongdoing", so presumably Moore would expect members of the legislature to *use* said theory of legislative aim.

institutional context. It is not a question of what makes the conduct morally wrongful, but instead of what epistemic method will you need to employ in order to find the property of moral wrongfulness in the first place.

The epistemic question is also important for the declarative stage of criminalisation, with regards to how potential hearers of what is being said through criminalisation will understand the meaning of conduct being morally wrong. In other words, if we were in fact saying something about the moral wrongfulness of a criminalised conduct by criminalising it, how do people understand that meaning? This will be explored in more detail in chapter 3, but I will say at this point that if we do in fact report moral facts about the criminalised conduct through criminalising it, then it would seem plausible to think that one can gain access to the moral properties of the criminalised conduct through a criminal offence. That is, there would be potential for some kind of moral learning – of gaining epistemic access to moral facts – through understanding what is being said by criminalising a conduct. So, the epistemic question is useful to explore this issue, specifically regarding the potential of epistemic access to moral knowledge through engaging in moral talk.

Therefore, the focus of the epistemic question is slightly different from that of the conceptual and ontological questions. The point of asking this question is less about the nature of moral properties and more about *how to find* moral properties in an epistemically sound way. It is, in this sense, a practical question about how to do things with moral properties and, more specifically, what the methodology to achieve this entails for how we should go about looking for the relevant wrongfulness to criminalisation.

2.3.1. The Possibility of Moral Knowledge

A central issue to the epistemic question, as mentioned above, has to do with the possibility (or lack thereof) for moral knowledge to exist. Specifically, the question I want to pose is whether morality is the kind of thing that one can gain *epistemic access* to. That is, is morality the kind of thing that can both fulfil the requirements for something to count as knowledge,[122] and that we can put ourselves in a position, as agents that are capable of acquiring knowledge, that allows for us to gain access to said knowledge. A full answer to the issues I discuss here needs to explain how moral knowledge exists (and if it does in the first place) and, if so, is it the kind of knowledge we can gain access to and how do we do so.

As one might imagine, and as I mentioned at the beginning of this chapter, a lot of the groundwork for answering the epistemic question will depend on our answers to the ontological question. A clear example of this is a strong anti-realist response to the ontological question, such as moral error theory. If we hold the belief that moral properties are *never* instantiated in reality, and that therefore all non-negative moral utterances are untrue, then we immediately hold a sceptical position about moral knowledge – there cannot be such a thing as moral knowledge because one cannot hold a moral judgment that is *true*. So, the answer to the epistemic question I pose here from the moral error theorist's view, would look something like the following: we cannot gain epistemic access to moral knowledge because there is simply no moral knowledge to access.

[122] Traditionally, a true, justified belief, though this orthodoxy has been challenged in recent years. Famously, see Edmund L Gettier, 'Is Justified True Belief Knowledge?' (1963) 23 Analysis 121.

In contrast, for a robust moral realist – who, as a reminder, holds that there are mind-independent moral properties to which moral language can refer and that are sometimes true – the groundwork to answer the epistemic question looks very different. To start, the robust realist accepts the possibility of there being moral knowledge, and that one can put oneself in an epistemic position to access said knowledge. The key, however, is how the robust realist believes one does this latter task – acquiring moral knowledge is a matter of accessing mind-independent properties in the actual world, so the epistemic position that will need to be taken is one that is effective in gaining access to these kinds of properties. Again, the answers given to the ontological question will affect what we might think this entails. If the robust realist in question is a naturalist, then the epistemic position to take is one where putative natural moral properties can be observed, measured or perceived in some way and processed into moral knowledge. If the robust realist is a non-naturalist, then the epistemic position to take is one where putative non-natural properties can be obtained, for example through some kind of rational deliberation process that does not depend on reductions to natural properties. That being the case, we can see that to determine the contents of the answer to the epistemic question, much rides on the answers one has already committed to for the ontological question.

These differences are important to tease out for many reasons. One of them is the fact that it can help us better understand how potential moral disagreements (assuming that genuine disagreement is possible) works with regards to *how a moral judgment was obtained*. Let us assume for a moment that there were genuine moral facts about the world and actions (thereby setting aside strong anti-realism). If this is the case, then differences between moral judgments will be based in large part on differences in *evidence* about the moral facts on which the judgment is grounded. In other words, one

important reason why we would believe that someone is mistaken about their moral view or judgment is that they have *misunderstood* or just *missed* evidence to the contrary of their view. If that is so, then a good way to get through the impasse of the disagreement is to look at the evidence being used as the basis for the moral judgments in question and compare how each of the disagreeing parties have obtained, interpreted, and analysed that available evidence. It will be particularly important when doing this to have a clear standard for what kind of evidence for moral judgments is appropriate, which in large part will be dependent, again, on where we believe one is supposed to go to in order to access those facts. Because note that if one is committed to the belief in there being genuine moral facts, then that entails the possibility of getting those facts wrong in an epistemic sense, i.e. making a mistake in our *knowledge* of those facts.

Of course, if we bring strong anti-realism back in and the moral scepticism it can entail, then the problem is even greater. Not only does a moral sceptic think there is no such thing as moral knowledge, but that also entails that there is nothing to genuinely disagree about here – both parties in the putative disagreement are simply mistaken about the matter-of-factness of what they are talking about. There are no facts – at least, not in the same sense as things like scientific facts about the physical world – on which to disagree and, therefore, have someone 'get it right' and another 'get it wrong' in their deliberations. As Wong explains it with regards to Mackie's view, "whereas disagreements over the physical structure of the universe are best explained by appeal to a lack of definitive evidence and speculation from the available evidence, moral disagreements are best explained as originating from differences in the ways people have chosen to live rather than their attempts to track independently existing moral facts or properties that make their ways

of life good or right."[123] Even if one were to disagree with Mackie's original idea of moral disagreements being based on "people's adherence to and participation in different ways of life"[124] as the alternative to being akin to facts about the physical world, the underlying point can still stand – moral facts might not be the kind of facts for which one can provide evidence in the same way as other kinds of facts. And if this is so, then *where* those putative moral facts come from, in epistemic terms, becomes a key point to figure out when trying to deal with potential disagreements on moral judgments.

The contrast need not be, however, only with scientific facts about the physical world – there are other kinds of facts that are not facts about the physical world that one might believe are a part of a plausible ontology, such as analytic facts, facts about fictional worlds, conceptual facts and so on. And even then, one might still think that those kinds of facts are different to moral facts. For example, one may think that there is such a thing as *legal* facts – facts of the matter on what the law prescribes – and believe that they are not the same as moral facts.[125] In this sense, one might believe that a disagreement about legal facts can be resolved – similarly to scientific facts – by appealing to an account of how we know those facts, and still hold that this is not the case for moral facts. Or contrastingly, one might believe that legal and moral facts work the same way epistemically, and so both kinds of potential disagreements can be resolved by observing the available evidence to which we have access.

[123] David B Wong, 'Relativism and Pluralism in Moral Epistemology' in Aaron Zimmerman, Karen Jones and Mark Timmons (eds), *The Routledge Handbook of Moral Epistemology* 317.
[124] Mackie (n 96) 36.
[125] For a discussion of this topic, see Brian Leiter (ed), *Objectivity in Law and Morals* (Cambridge University Press 2001).

It is also at the stage of the epistemic question where the issue of deciding whether moral properties are 'discovered' or 'constructed' becomes relevant: which view is accepted will determine how the epistemic method ends up looking like. To put it differently, let us imagine that we are faced with a particular conduct, say meat-eating. When trying to determine whether meat-eating is morally right or wrong, there are basically two ways to go about it. We can either try to find moral wrongfulness within that conduct by some kind of cognitive process for obtaining knowledge from the evidence we have available, and then show our thinking as to how we have found the property of moral wrongfulness in meat-eating. Conversely, we can go the opposite direction and impose or ascribe moral wrongfulness on to meat-eating by some kind of cognitive process that creates or constructs moral knowledge and then show our thinking as to how we have determined that moral wrongfulness is aptly ascribed to meat-eating. The first process is of discovering moral wrongfulness in meat-eating, the second is of constructing moral wrongfulness from meat-eating. And both processes require different methods to gain epistemic access to the facts that will lead to either discover or construct moral wrongfulness as a property of meat-eating.

Ultimately, however, the important point for my purposes is that if we accept the view that moral knowledge is possible, then we must then continue to ask ourselves *how* that knowledge is acquired. In other words, how does epistemology work with moral properties? This is what I turn to in the next section.

2.3.2. Moral Epistemology and Epistemic Method

Moral epistemology is, as the name suggests, the theorising of how moral knowledge works. In the discussions I had above, I have already covered

some of the broad issues about moral properties that are relevant to moral epistemology more generally, such as whether moral properties can be seen as facts or not in the first place. However, for this section I will focus on one important epistemic issue that will be relevant to criminalisation theorising: the issue of epistemic justification through pursuing a particular epistemic method. When trying to justify a belief from an epistemic perspective, we need to look at how we have inferred said belief from the evidence that we have available. A familiar problem with this is the 'epistemic regress argument',[126] by which one can seemingly trace back beliefs that justify our inferred beliefs *ad infinitum*. In the context that I am interested in here, the epistemic method that one decides to use to justify the belief in a moral property will affect how one deals with this potential regress problem. One way to deal with this, as mentioned before, is to simply accept scepticism and the fact that our moral beliefs are never justified epistemically.[127] However, if a theorist does want to accept that there can be justified moral beliefs, then there are a few ways in which this can be argued for, based on the epistemic method that one chooses.

A first option is moral intuitionism, which entails the claim that "some people are justified in believing some moral claims in some way that does not depend on the believer inferring or even being able to infer the claim from anything else that the person believes."[128] In other words, the intuitionist will commit to the idea that the regress for justification stops at some point, and it stops at the level of beliefs that are non-inferentially justified. Importantly,

---

[126] Walter Sinnott-Armstrong, *Moral Skepticisms* (Oxford University Press 2006) 74–77; Andrew Fisher, *Metaethics: An Introduction* (Taylor & Francis Group 2014) 142 <http://ebookcentral.proquest.com/lib/ed/detail.action?docID=1886850> accessed 29 June 2023.

[127] See Matt Lutz and Jacob Ross, 'Moral Skepticism' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2018).

[128] Railton (n 79) 25.

these intuitive moral beliefs, by not needing to be inferred, are self-justifying. Note, however, that this self-justification does not mean that the intuitionist cannot reflect on this non-inferential belief in order to *understand* it, but rather it only means that, as Sinnot-Armstrong explains, "if moral intuitionists require some kind of reflection, such reflection must help only to understand the belief and must not provide premises for inferences to the belief."[129] One of the advantages of this view is that, as Tropman puts it, "it is true to our moral experiences. We seem to arrive at moral judgements, both singular and general, immediately and without inference."[130] However, there are several problems with this epistemic method. One is the point shown earlier that Mackie makes with his argument from queerness, where our moral intuitions are accessing a mysterious non-natural property, for which we also have a mysterious faculty to intuitively access those properties.[131] Further, as Tropman tells us, "for the intuitionist, the target moral beliefs would be indubitable, obviously true, and infallible, all of which flies in the face of the complexity and uncertainty of moral practice."[132] So, moral intuitionism might give us some initial clues as to the presence of moral properties, but it is possible to argue, as Sinnot-Armstrong warns, that "moral intuitionism cannot show how any moral believer or belief can be justified in any way that is sufficient to stop the skeptical regress."[133] There are possible responses, however, to some of the criticisms made to intuitionism and it is still a plausible epistemic position to hold.[134]

---

[129] Walter Sinnott-Armstrong, 'Moral Skepticisms and Justification' in Walter Sinnott-Armstrong and Mark Timmons (eds), *Moral Knowledge? New Readings in Moral Epistemology* (Oxford University Press 1996) 26.
[130] Elizabeth Tropman, 'Intuitionism in Moral Epistemology' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2018) 473.
[131] Mackie (n 96) 39.
[132] Tropman (n 130) 474.
[133] Sinnott-Armstrong, *Moral Skepticisms* (n 126) 219.
[134] For a robust defence, see Robert Audi, 'Moderate Intuitionism and the Epistemology of Moral Judgment' (1998) 1 Ethical Theory and Moral Practice 15.

A second option is what is called moral coherentism. This view states that "a belief is justified when it has the right relation to a system of beliefs that have the right relation to each other",[135] and that 'right' relation is coherence. The main exponent in moral and political philosophy of this method was John Rawls,[136] with his idea of 'reflective equilibrium'. In what is probably the starting point of writings discussing coherentism in moral epistemology, Sayre-McCord characterises Rawls' method "as a matter of shifting back and forth among the various moral judgments one is initially inclined to make and the more or less abstract theoretical principles on is examining and attempting to develop",[137] which leads to a constant revising and evaluating of one's beliefs and to the development of a *system* of beliefs, that end up creating a basis for further inferences. This may, at first, read as a circular kind of epistemic justification. However, as Fisher points out, the coherentist is using a holistic sense of justification, which is best explained through the metaphor of a raft:

> "A raft stays afloat not because of a particular plank (belief) but because of how the whole structure hangs together (whole set of coherent beliefs). Because of this holistic account of justification the coherentist can talk about a belief being justified without the charge of circularity. The fundamental unit of evaluation when it comes to epistemic justification is the whole of the system of beliefs an agent has."[138]

It is easy to see the attractiveness of this kind of epistemic method, since it allows for a systematisation of beliefs into a stronger sense of justification, that is based on requirements such as logical consistency, evidential

---

[135] Sinnott-Armstrong, *Moral Skepticisms* (n 126) 220.
[136] Rawls (n 111).
[137] Geoffrey Sayre-McCord, 'Coherentist Epistemology and Moral Theory' in Walter Sinnot-Armstrong (ed), *Moral Knowledge? New Readings in Moral Epistemology* (Oxford University Press 1996) 141.
[138] Fisher (n 126) 151.

consistency and relevance.[139] However, coherentism is not free of criticism. First, it is quite possible that nobody actually has a consistent, coherent system of moral beliefs, which requires further empirical investigation. Second, even if there were people who do have coherent systems of beliefs, it is also quite possible that people have *different* sets of coherent systems of beliefs, and if coherentism is true then all of those people can justifiably deny the system of beliefs of the others.[140] Finally, the coherentist account does not mention a requirement for *truth* of the system of beliefs, so this leaves the door open for a situation where, as Fisher tells us, "if our belief that killing is wrong is justified because it is part of the most *coherent* set of beliefs then this does not meant that the belief that killing is wrong is *more likely* to be true."[141] Again, however, there are further answers to these criticisms, and coherentism remains a plausible option as a moral epistemic method.[142]

Finally, I will mention the option of contractarianism. As an epistemic method, contractarianism states that "a person can be justified in holding a moral belief if he or she is justified in believing that the moral belief is justified by some agreement."[143] Some prominent examples of this kind of epistemic method are the works of Scanlon,[144] Habermas,[145] and Gauthier.[146] The basic idea of this kind of epistemic method is that the act of agreeing to the contents of a moral belief by an appropriate set of agents leads to epistemic justification of that moral belief. Whether it is from a position of universal rational agreement,[147] or from a position from which no one could reasonably

[139] See Sayre-McCord (n 137) 166–170.
[140] Sinnott-Armstrong, 'Moral Skepticisms and Justification' (n 129) 32.
[141] Fisher (n 126) 153.
[142] For some of those replies, see Sayre-McCord (n 137) 170–177.
[143] Sinnott-Armstrong, 'Moral Skepticisms and Justification' (n 129) 37.
[144] Scanlon (n 112).
[145] Jürgen Habermas, *Moral Consciousness and Communicative Action* (Polity Press 1990).
[146] David Gauthier, *Morals by Agreement* (Oxford University Press 1987).
[147] ibid.

reject a set of principles,[148] the common point is that this contract, this agreement is what can provide justification for the belief in certain moral properties such as moral wrongness. This kind of epistemic method is particularly attractive to theorists that may want to emphasise the liberal commitments of political systems, but it is also not free of criticism. For a start, most of these approaches use some idealised conceptualisation of rationality, reasonableness, or discourse which, for example, "have come under frequent attack by feminist philosophers for ignoring the differences between people and thus overestimating our ability to reliably predict agreement."[149] There are also criticisms regarding determining membership to the original agreement, and in determining the different moral standings of members of the agreement, where contractarians do not provide enough in order to establish the proper scope of those who enter the agreement.[150]

The point of giving an overview of these different possible methods for moral epistemology is not to provide an in-depth analysis of each one, nor to advocate for any of them as the right choice. Instead, the point is to precisely show that *all* of these options are plausible as epistemic methods, but that they can lead to very different understandings of how a moral property like moral wrongness is supposed to be found. And, as we shall see in detail in chapter 4, this is especially important for the theorising of criminalisation when it comes to our normative theorising of how a decision on criminalising a conduct should be done.

---

[148] Scanlon (n 112) 153 "An act is morally wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulations of behavior that no one could reasonably reject as a basis for unforced general agreement".

[149] Amia Srinivasan, 'Feminism and Metaethics' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2018) 604.

[150] Christopher W Morris, 'A Contractarian Account of Moral Justification' in Walter Sinnot-Armstrong and Mark Timmons (eds), *Moral Knowledge? New Readings in Moral Epistemology* (Oxford University Press 1996) 222–223.

So, in conclusion, there are two important issues that moral epistemology can help us understand about the deliberation process that a potential criminalisation decision-maker is going to undertake: *how* one gains knowledge about moral wrongness, and what the *target* of that knowledge is, or put differently, what someone is looking for to gain that knowledge. The important thing to notice, however, is that there are *different* potential answers to those methodological questions, and those answers are dependent on broader metaethical assumptions and commitments that one may have in trying to answer those questions. The specific questions of how moral epistemology's methodology affects criminalisation theories requires understanding the assumptions being made to answer the ontological and conceptual nature of the moral facts being sought after, in this case moral wrongness. This will be particularly salient regarding the target of the inquiry, because in order to properly define a methodology for the deliberation, we need to know the nature of what it is that we are looking for. As Sinnot-Armstrong tells us, "we cannot determine whether any claim is justified or known if we have no idea what that claim means, so moral epistemology depends in some ways on moral semantics. […] The same might be said for moral ontology, the definition of morality, deontic logic, and moral psychology."[151] Not having this clarity, as Fumerton puts it, "would be like going on safari to hunt a rhinoceros without knowing what a rhinoceros is."[152] Knowing what our 'rhinoceros' is will not only inform the target of the inquiry, but will also affect our *evaluation* of the methodology being used. To continue Fumerton's metaphor, once we know what a rhinoceros is we can then determine whether it is better to hunt one with a shotgun or a slingshot, whether we should hunt them at night or during the day, and whether we should go hunting by ourselves so as not to be detected or as a group to have

[151] Sinnott-Armstrong, *Moral Skepticisms* (n 126) 7.
[152] Richard A Fumerton, *Reason and Morality* (Cornell University Press 1990) 4.

safety in numbers – or even if we should go hunting at all.

This last issue of evaluating epistemic methodology is what is most relevant to my endeavours in this thesis. In spite of the fact that most criminalisation theorists are not explicit in providing their views of how the seeking out of moral wrongness is supposed to be done, I will show in chapter 4 that one can construct an idea of what this methodology might be (i.e. how a theorist might answer the epistemic question) from the metaethical assumptions one can infer from their work. In doing this, my aim will be to bring attention to a particular kind of evaluation that one can make of criminalisation theories which use moral wrongness as part of the deliberation process: whether there is an appropriate way by which a decision-maker can gain epistemic access to the moral wrongness of conduct. In this context, 'appropriate' will mean different things, depending on the metaethical commitments that one has about moral facts. This reinforces the need to be as clear as possible about any metaethical assumptions one is making in theorising about a deliberation process for criminalisation.

In conclusion, this chapter has posed the three main metaethical issues or questions – the conceptual, ontological, and epistemic questions – that this thesis will be exploring regarding the normative theorising of criminalisation, as well as a brief summary of different metaethical stances and views surrounding those questions. Once again, the purpose of this chapter has not been to advocate for any of these views, but instead my point is to provide some basic understandings of what metaethics is *about*, and how the discussions present in metaethics can inform our theoretical understanding of criminalisation. How this is the case specifically, will be the focus of the next chapters of the thesis.

CHAPTER 3 – THE DECLARATIVE STAGE OF CRIMINALISATION

In this chapter, I will claim that when a conduct is criminalised, something is being *said*. That is, when formally defining a conduct as a criminal offence – usually by the legislature but it does not necessarily have to be them – communication seems to be happening towards people within a jurisdiction. This goes beyond just the fact that criminalising a conduct requires the use of language – as any act done within the law more generally also does – but rather the point is that in the act of speaking[153] the words by which a conduct is criminalised, we are also *performing* further actions that are not just the uttering of words of a particular language.

For example, if someone says to you "could you pass the salt, please?" they are not only asking you a question about your capacity to pass salt to them – in fact, it would be very strange to understand that utterance as that kind of question. Instead, you would interpret it as a *request*, as something that is more akin to them saying "I want salt, there is salt close to you and I want you to pass it to me", though of course it seems improbable that someone would ever say it in those terms. And yet, that is precisely what is being conveyed when they say "could you pass the salt, please?" In saying this, they are *acting* – they are making a request and, incidentally, they are also letting you know that this is being asked according to the norms of etiquette, since they are asking "please". They are letting you know that it is a polite request, rather than a command or an order. And, importantly, all of this is happening even when they are not *literally* saying any of this to you – the utterance of the words *means* that they are doing what I have described earlier, and both a competent speaker and hearer can understand this

[153] 'Speaking' here does not need to be read literally – a piece of writing, for example, can also 'speak' in the sense I am using it here.

meaning. This is what in the philosophy of language is called a speech-act, and I will propose that the act by which we criminalise a conduct through the criminal law is, indeed, a speech-act. That is, criminalisation can be seen as an act that is performed by saying that one is doing so, as long as one has the right authority to do said speech-act. Again, this does not need to be done literally – it is enough that actions are performed (i.e. speech-acts) that carry with them the meaning of criminalising a conduct.[154] For the purposes of this thesis, I will refer to this speech-act as a *criminalisation claim*, which I understand as a speech-act the performance of which counts as the official act by which a conduct is, from the moment of utterance onwards, considered a criminal offence.

The central issue for the declarative stage of analysis, then, is defining exactly what is being said when a conduct is criminalised. To answer this, I will be using some basic tools from speech-act theory to analyse the potential contents of a criminalisation claim. The purpose of doing this is to propose a model for understanding the act of criminalisation as several normatively relevant things happening at once, all of which are done through the act which makes a conduct a criminal offence.

Before doing so, a quick word on why it is useful to look at criminalisation as a speech-act. To begin, it seems to be a consensus among criminal law theory scholars that when a conduct is criminalised, some form of communication is happening. Simester and Von Hirsch propose that criminalisation "speaks directly to subject-citizens"[155]; Tadros believes that it

[154] Though, interestingly, many legislatures include texts to the effect of uttering the speech act in question, such as the text included in the Acts of Parliament for the United Kingdom, which include the text "Be it enacted by the King's most Excellent Majesty, by and with the advice and consent of the Lords Spiritual and Temporal, and Commons, in this present Parliament assembled, and by the authority of the same, as follows:-"
[155] Simester and Von Hirsch (n 1) 12.

"expresses to citizens that they are under obligations" and that the criminal law "communicates moral duties to citizens";[156] Duff states that the criminal law "should be seen as a communicative enterprise".[157] Husak, echoing Feinberg,[158] speaks of the expressive function of the criminal law and that "an important objective of the criminal law is to convey both to the offender and to the community the wrongfulness of his conduct."[159] It is not surprising that theorists propose this kind of analysis of the criminal law and particularly of criminalisation, since creating a criminal offence – and thereby bringing the criminalised conduct into the purview of the criminal law – seems to carry with it many potential *meanings*.[160] This 'carrying' of meaning – the idea that by saying something we are also making it the case that other things happen – is best seen, I propose, as a kind of speech-act.

The chapter continues by setting out what speech-act theory is, focusing specifically on some of the key terminology used in this kind of theorising, such as illocutions, background information and Searle's taxonomy of illocutionary acts. Importantly, I will follow Searle's distinction between 'assertions' and 'declarations' to identify the difference between reporting normative facts and creating new normative facts by declaring them into existence. With this distinction in mind, I then propose a theoretical model for criminalisation as a speech-act. In it, I propose that the act of criminalisation can be seen as an indirect speech-act, which I call a criminalisation claim, that both declares into existence new normative facts relative to the conduct being criminalised, as well as asserting the normative status of that conduct.

---

[156] Tadros, *Wrongs and Crimes* (n 1) 159–160.
[157] Duff, *The Realm of Criminal Law* (n 1) 109.
[158] Joel Feinberg, *Doing and Deserving* (Princeton University Press 1970).
[159] Husak, *Overcriminalization: The Limits of the Law* (n 1) 139.
[160] Regarding the assigning of meaning to criminalisation and its potential symbolic uses, see Javier Wilenmann, 'Framing Meaning through Criminalization: A Test for the Theory of Criminalization' (2019) 22 New Criminal Law Review.

In doing so, the speech-act makes it the case that the conduct is now a criminal offence. I argue that the declared normative facts are the labelling of the conduct as an offence, the creation of an obligation not to do the conduct, the creation of a liability to punishment for doing the conduct, and the creation of a responsibility to answer for when one carries out the conduct. Importantly, I argue that the normative status of the conduct, which I propose is wrongfulness in a broad sense, is being made implicitly in the speech-act, and so that assertion needs to be inferred by hearers of the speech-act from the context of utterance. I show that one way to argue for how this happens is through the declared normative facts that are created by the speech-act acting as the context by which a hearer can infer that, if those normative facts are being declared, then the conduct is probably being portrayed as wrongful.

I then argue that, if one wants to propose a theory that uses moral wrongfulness as part of the explanation for what is being said through criminalisation, that moral wrongfulness is best seen as being included in that implicit assertion of wrongfulness. From there, I bring in the metaethical questions to show how our understandings of the conceptual, ontological and epistemic nature of moral wrongness can affect our theoretical understanding of how the inference of moral wrongness from the declared normative facts works. I end the chapter by suggesting that these metaethical issues need to be addressed by theorists in order to make sense of the fact that a hearer can reasonably infer that the wrongfulness being implicitly asserted is specifically of a moral kind.

3.1 Basic Thesis and Speech-Act Theory Terminology

So, what are the contents of a criminalisation claim? To answer this, I will

need to explain some basic concepts of speech-act theory, but here is the basic thesis that I will be defending: a criminalisation claim is an indirect speech-act which both asserts and declares normative facts[161] and, through the utterance of the speech-act, a target conduct becomes a criminal offence. I will explain the terminology used within that basic thesis and then look at the actual contents of criminalisation claims.

First, a quick explanation of what 'illocution' means is necessary, since understanding the rest of the terms in the basic thesis requires knowing this. The term comes originally from Austin,[162] and refers to what is being *done* in saying something (i.e. the speech-act itself) rather than just the literal meaning of the words being used (in Austin's terms, locutions). Some classic examples of illocutions are common verbs like requesting, questioning, bargaining, promising, commanding, and so on – if I say the phrase "where is the library?" the illocution is a question, which is being performed *by uttering* the phrase "where is the library?"

Next, indirect speech-acts are, as Searle puts it, when "one kind of illocutionary act can be uttered to perform, *in addition*, another type of illocutionary act."[163] The example of a speech-act I gave earlier – saying 'could you pass the salt, please?' – is a good example of this kind of speech-act (and is based on a similar example by Searle).[164] One meaning of the speech-act is a mere question: the speaker is asking whether the hearer is indeed capable of passing the salt. But as discussed earlier, it seems quite natural to

[161] By 'normative facts' here, I mean states of affairs that are, in some important sense, normative. They *can* be seen as moral facts, but at this stage I wish to leave the door open for those facts to be considered as part of any potential normative domain.
[162] See JL Austin, *How To Do Things With Words* (Oxford University Press 1975).
[163] John Searle, *Expression and Meaning - Studies in the Theory of Speech Acts* (Cambridge University Press 1979) 30. Emphasis in the original.
[164] John Searle, *Speech Acts - An Essay in the Philosophy of Language* (Cambridge University Press 1969).

understand the meaning of this utterance as not just a question, but a *request*: the speaker is asking the hearer to pass the salt to them. The utterance of 'could you pass the salt, please?' also carries with it a kind of 'marker' of the nature of the request: by asking 'please', the speaker lets the hearer know that they are trying to follow the norms of etiquette by asking politely. As you can see, all of these meanings are happening *at the same time* as the utterance 'could you pass the salt, please?'. You will also notice that these additional meanings are *not* literally uttered in the speech-act – they are obtained from what Searle calls "mutually shared background information"[165] and from the capacity of the hearer to make inferences based on that background information.

Seeing a criminalisation claim as an indirect speech-act makes sense when we actually look at how criminal offences are worded and then compare that to how many theorists understand what is happening when a conduct (Φ) is criminalised. For example, if any theorist wants to suggest that criminal offences convey some kind of wrongfulness, then the only option for this to be the case is to understand criminal offences as indirect speech acts, since there is no offence that literally says anything like "it is morally wrong to Φ", "citizens shall not do Φ" or "it is prohibited to Φ". For example, s. 47(1) of the Criminal Law (Consolidation) (Scotland) Act 1995 states that "Any person who has with him in any public place any offensive weapon shall be guilty of an offence" – nowhere in the statute does it say explicitly that the relevant conduct is wrongful. It only states that performing the conduct makes one 'guilty of an offence' and then sets a liability for doing so.[166] One might think

[165] Searle (n 163) 31–32.

[166] s. 47(1) continues "[…] and shall be liable-
(a) on summary conviction, to imprisonment for a term not exceeding six months or a fine not exceeding the statutory maximum, or both;
(b) on conviction on indictment, to imprisonment for a term not exceeding 5 years or a fine, or both.

this is the case because carrying an offensive weapon is a *mala prohibita* offence, but offences traditionally seen as *mala in se* are also expressed in a similar fashion. The offence of rape is set out in the Sexual Offences (Scotland) Act 2009 s.1, which states:

"(1) If a person ("A), with A's penis-

(a) without another person ("B") consenting, and

(b) without any reasonable belief that B consents,

penetrates to any extent, either intending to do so or reckless as to whether there is penetration, the vagina, anus or mouth of B then A commits an offence, to be known as the offence of rape."

Once again, there is no explicit signalling of the wrongfulness of this conduct in the wording of the offence, nor of the fact that the conduct is prohibited – it describes the relevant conduct and assigns to it the label of an offence, specifically the label of rape. And yet, it does not seem implausible to believe that the wrongfulness of the conduct at play is being conveyed through criminalising it – there is something in the fact that it is being criminalised that generates this further or additional meaning of wrongfulness and I argue that the best way of understanding this is that it is an additional illocution that is being performed at the same time of the utterance of the act which criminalises Φing.

This is because indirect speech acts rely on a shared background information which allows for the hearer to make particular rational inferences from what is being said[167], and this seems to fit with what is actually happening with conduct being criminalised, or at least it fits with how many theorists understand what is happening. As stated earlier, the fact that a conduct is included into the criminal law as an offence seems to 'carry' with it

[167] See Searle (n 163) 31–32.

more than just one meaning, and this 'carrying' is best seen as additional illocutions happening at the same time. Included in this background information are things like knowing about the criminal justice system, the existence of criminal punishment and, more generally, the different kinds of norms (from different normative domains) that may have something to say about the conduct in question. For example, you will notice that an offence is not worded literally in terms such as 'the killing of another is morally wrong', and yet it does not seem an irrational inference to understand, from the fact that murder is criminalised, that the law is saying that it is morally wrong to kill another. This inference is possible in virtue of the shared background information which exists between speaker and hearer. I will explore this issue of inferences and the context required to do so in more detail in the final section of this chapter, but for now it is enough to say that indirect speech acts require this kind of inference process, because the literal meaning of the indirect speech act does not necessarily match the meaning of the illocution being performed. I will use an example based on Searle's work[168] to illustrate this point.

Suppose that we have two people, Amy and Beatrice, that have the following exchange:

(1) Amy: Let's go to the cinema tonight.

(2) Beatrice: I have to study for an exam.

(2) is an indirect speech act performed by Beatrice because, using Searle's terminology, the *primary* illocutionary act performed by Beatrice is a rejection of an invitation, and the *secondary* illocutionary act is an assertion of the fact that Beatrice has an exam that she has to study for.[169] Notice that the

[168] ibid 33–36.
[169] ibid 33.

secondary illocution Beatrice performs is done *in the act of saying* what she literally says – she says that she has to study for an exam. The primary illocution, however, is not literal in the same way, because nothing in the words that Beatrice uses directly states that she is rejecting Amy's invitation. In fact, she could use the same words in a sentence that cancels out the primary illocution of rejecting an invitation if she added more words, such as: "I have to study for an exam, but I'll do it when we get back from the cinema". And yet, it seems quite plausible that Amy (and us) can infer, if Beatrice says nothing more than what she says at (2) as a response to (1), that she is rejecting the invitation to the cinema. How does this happen? Searle tells us that a hearer of (2) needs to have an inferential strategy in order to realise that an indirect speech act is happening, and he describes it thusly: "the inferential strategy is to establish, first, that the primary illocutionary point departs from the literal, and, second, what the primary illocutionary point is."[170] In addition to this inferential strategy, a hearer will also need "a device for finding out what the ulterior illocutionary point is",[171] which is "derived from the theory of speech acts together with background information."[172] Thus, for the inference to work, a hearer needs to be able to identify a discrepancy between the literal meaning of the utterance and the potential illocution being performed in the speech-act. Once this is done, they can move on to determine what that illocutionary point is, based on their understanding of the background information that applies to the communication taking place.

Finally, Searle proposes a taxonomy of categories for possible illocutionary acts,[173] where he proposes five kinds of illocutions: assertives,

[170] ibid 35.
[171] ibid 47.
[172] ibid 48.
[173] See Chapter 1 in Searle (n 163)..

directives, commissives, expressives and declaratives (sometimes called 'exercitives' but I will use Searle's original terminology for simplicity). Of these five, I wish to focus on two: assertives and declaratives. Assertives are illocutions that commit the speaker to something being the case (i.e. the truth value of what is being said), whereas declaratives are illocutions the successful performance of which guarantees that the propositional content of what is being said corresponds to the world. In other words, assertions are reports of states of affairs which the speaker believes to be true, whereas declarations are illocutions which create new states of affairs through their performance. For example, when a meteorologist utters the phrase "it is raining outside", it is an assertion – it is a report of how things actually are and, if the speaker is being sincere (i.e. they are not knowingly saying something false), they actually believe the report to be true (which we can corroborate by looking outside, so the assertion is also falsifiable). In contrast, when the city official utters the phrase "I hereby pronounce you husband and wife", they are making a declaration – by uttering that phrase, they are inserting a new fact into the world and, if they are successful (i.e. they have the required authority to do so) then that fact comes to be true in the act of uttering it. There is a potential further sub-division of declaratives, proposed by Bach and Harnish,[174] of 'effectives' and 'verdictives', where the former are utterances that when issued by the right person under the right circumstances produce institutional states of affairs, whereas the latter are reports of something being the case and, in doing so, make that the official line.[175] However, I will prefer Searle's terminology since I wish to remain agnostic as to which specific type of declarative illocutions are happening in a criminalisation claim – choosing between effectives and verdictives may be based on metaethical assumptions of the kind of normative facts that are

[174] Kent Bach and Robert Harnish, *Linguistic Communication and Speech Acts* (MIT 1979).
[175] See Nicholas Allott and Benjamin Shaer, 'The Illocutionary Force of Laws' (2018) 61 Inquiry 351, 354.

being declared in criminalisation claims, so I will not impose these assumptions on the framework I am proposing here.

3.2. Declared Normative Facts

Now, let us shift our attention back to criminalisation as a speech-act. As I said above, a criminalisation claim is an indirect speech-act that both asserts and declares normative facts, and makes it the case that a conduct becomes a criminal offence. Since I am arguing that it is an indirect speech-act, it has both a primary and a secondary illocutionary act happening at the same time, and the primary illocution is not going to be explicitly stated in the speech-act. What *is* stated, however, are the effects of a conduct becoming a criminal offence.[176] These are, I argue, the normative facts that are being *declared* through the speech-act, in Searle's terms. Importantly, it is from these facts that are being declared that a hearer would need to be able to tell, from the speech-act and the background information or context of utterance, that an additional, primary illocution is happening at the same time.

As for what is being declared by a criminalisation claim – what new facts are being created in criminalising Φ – I propose that there is one descriptive fact and three legally relevant normative facts. The first and more obvious descriptive one I call the *labelling* fact, which is the assigning to a particular legal category (i.e. a particular criminal offence) of Φing. We declare that the factual elements that compose Φ (i.e. the elements of the offence) are hereby labelled as the particular criminal offence that is being established by the criminalisation claim. So, for example, if Φ were an attack on another person with some relevant intention (depending on the jurisdiction in question), by

[176] Granted, some of these statements are also not necessarily explicit in the speech-act, but they can be attributed to the speech-act in view of the fact that it is being spoken through the criminal law, which I will explain briefly for each case where this happens.

criminalising Φ we are declaring the assigning of the label "assault" to the elements which constitute Φing. Sometimes, however, this labelling fact is not necessarily present in the act of criminalisation, in the sense that the offence being created is not given a particular name such as 'assault' or 'culpable homicide', but is rather *ex post* named simply by its location in legislation – 'the offence under s. 47(1) of the Criminal Law (Consolidation) (Scotland) Act', to repeat a previous example.

Additionally to the labelling fact, there are three other new normative facts being declared. I will call them the *obligation* not to Φ, the *liability* to punishment for Φing, and the *reponsibility* for Φing. The first normative fact I will explain is the new liability when someone Φs, specifically linked to a breach in the previously declared obligation not to Φ. By criminalising Φ, we are letting people know that it is hereby declared that whoever Φs has breached a legal obligation not to Φ and, therefore, is liable to face a particular consequence. Normally, this consequence would be some form of criminal punishment, and though this may not necessarily need to be something like imprisonment, it will still be condemnatory in nature.[177] Again, this liability did not exist, as a matter of fact, until it was declared by the act of criminalising Φ – there may have been other kinds of consequences associated to Φing, but not the ones that result from the criminal law. This new liability is particularly important for those who attribute the distinctiveness of criminalisation to the kind of consequence people might face for performing the criminalised conduct, specifically when it comes to criminal punishment – in spite of any kind of other sanctions that may be associated to Φing (pre-legal or not), the act of criminalising creates a new, specific liability for the kind of sanction that is only available to criminal

[177] See Sec. II in Cornford (n 15). for a description of some potentially different legal liabilities that are condemnatory, yet not imprisonment, such as a criminal record and the effects of sentencing.

offences. Husak, for example, states that "the most basic questions to be answered by a theory of criminalization is: *for what conduct may the state subject persons to punishment?*"[178] Moore also considers that a theory of punishment goes hand in hand with a theory of criminalisation,[179] so he would also be interested in this liability. Any search for justification for criminalisation that is based on the kind of sanction that is applied to offenders will be interested in what kind of liability is being declared and what kind of consequence people will be liable for. As stated earlier, this liability also includes the consequence of being convicted for the offence in question, which can have its own distinct consequences that are different from mere punishment – having a criminal record, going on particular special registries (like some sexual offenders), as well as all the potential stigmas that are associated with being convicted of a crime independent of something like a prison sentence.[180] So, not only is there a liability for the punishment itself, but also for all that comes with the conviction that may lead to that punishment.

The second normative fact is the fact that criminalising Φ creates a new legal obligation with regards to Φing. Specifically, it creates an obligation not to perform Φ in view of the fact that it is now a criminal offence. This normative fact does not need to be stated explicitly in the secondary illocution of a criminalisation claim, but it is possible to view it as a part of the speech-act because of creating a liability for punishment for Φing. In this sense, both of these normative facts are connected to each other, and happen

[178] Husak, *Overcriminalization: The Limits of the Law* (n 1) 82. Emphasis in the original.
[179] Michael S Moore, 'A Tale of Two Theories' (2009) 28 Criminal Justice Ethics 27, 36; See also generally Moore, *Placing Blame - A General Theory of the Criminal Law* (n 1).
[180] See Zachary Hoskins, 'Criminalization and the Collateral Consequences of Conviction' (2018) 12 Criminal Law and Philosophy 625. For a recent analysis of these potential consequences in the context of the USA, see Sec. 1 in Jeffrey M Brown, 'Collateral Legal Consequences of Criminal Convictions in a Society of Equals' [2020] Criminal Law and Philosophy <https://doi.org/10.1007/s11572-020-09544-7>.

simultaneously – creating a liability to punishment entails creating an obligation not to perform the act for which one is liable to be punished. This is not an evaluative claim, in the sense of saying that the act is being punished because it is *wrong* to not follow the obligation being created, but rather it is a claim of normative entailment for the liability to punishment to make sense. That is, since the speech-act is creating a legal liability to a consequence (like punishment), there needs to be a correlative obligation for which the liability is responding to. Hence, in declaring a liability to punishment, the speech-act is simultaneously declaring an obligation.

Notice that this declaration does not mean that this new, created obligation is the *only* obligation that exists normatively speaking with regards to Φing. However, it *does* mean that criminalising Φ actually creates a new obligation, a legally relevant one that is officially integrated into the criminal law of a jurisdiction and declares that, from now on, citizens are under a *legal* obligation not to Φ. In essence, this new fact by which a legal obligation is created establishes a new reason not to Φ, which can be added into a decision-making process regarding Φing along with any other pre-existing reasons that one may identify with regards to Φing, but there now is an additional reason being created simply by the fact that Φing has been criminalised. The obligation is specifically in the form of a prohibition – this is precisely what Simester and Von Hirsch are referring to with the idea of an 'instruction' given through criminalising.[181] The normative force of the 'should-not' that said instruction or prohibition carries with it is in a way reinforced by creating new normative facts about the performance of Φ, thereby making it part of the normative domain of the law. An obligation to not-Φ is created – it is declared that from now on, there is hereby an obligation, relevant to the criminal law, not to Φ.

[181] Simester and Von Hirsch (n 1) 12.

Tadros has a related view with regards to what is being done when criminalising, particularly for what he calls the "core cases of criminalisation" in which "the criminal law expresses to citizens that they are under obligations. […] it demands that citizens must not perform the act prohibited by the criminal law because those acts are seriously wrongful. The criminal law thus communicates moral duties to citizens."[182] Tadros' view, however, is not clear on whether these obligations are, in the terms used here, asserted or declared. On the one hand, he accepts that there might be prior moral demands on the conduct being criminalised, so the 'communication' of moral duties might be a kind of mere assertive reinforcement of that putative previous duty. On the other hand, it seems that if the expression happening when criminalising a conduct is a demand on citizens not to perform Φ, that expression is adding something new to the normative status of Φing, rather than just reminding hearers of a prior moral duty. Note that saying that there is an obligation not to Φ *because* it is morally wrongful does not necessarily imply that the obligation is a pre-existing moral one – moral wrongfulness can be acting as a justifying reason for the creation of a new kind of obligation (like a legal one), which seems more in line with what Tadros is proposing.

This distinction is important because it entails two different views of what the law does when using moral wrongfulness in the context of criminalisation. One option is that the law is mirroring morality and simply reporting back what is purported to be the moral matter of fact with regards to both the moral properties of the conduct and the obligations that are generated from that fact. The other option is that the law is only using the fact that moral wrongfulness is a property of the conduct in question as a basis for creating new normative facts, which can function independently of the moral

[182] Tadros, *Wrongs and Crimes* (n 1) 159–160.

obligations one might purport to associate with the moral wrongfulness of the conduct.

Duff sees criminalisation more like the first option mentioned above, because to him a criminalisation claim is 'declaring' (or asserting, in the language I have been using thus far) rather than 'prohibiting' the conduct in question, in view of the fact that "its role is not to make wrong what was not already wrong, but to declare that these pre-legal wrongs are public wrongs".[183] In this sense, criminalisation claims are not prohibitions because the authority by which we refrain from Φing is derived from the pre-legal moral wrongness (and thereby the moral duties which that would entail) rather than from some kind of respect for what the law says as an authoritative restraint on conduct.[184]

Perhaps a useful way to put it is to see the assertion/declaration distinction as the difference between what is being said and what is being done through a criminalisation claim.[185] For what is being *said*, Duff's position seems correct if we assume that the wrongfulness in question is pre-legal: the wrongfulness of the conduct being criminalised is not derived from the fact that it is prohibited, but instead we are asserting ('declaring' for Duff) its wrongfulness through criminalising the conduct – recognising the fact that it is wrong and, therefore, worthy of being included as a criminal offence. For what is being *done*, however, Duff's position needs to be adjusted: prior to the criminalisation of the conduct, there may have been moral (pre-legal) duties

[183] Duff, *Answering for Crime - Responsibility and Liability in the Criminal Law* (n 1) 86.
[184] ibid 85–86.
[185] Note that I am not proposing that when we say something we are not also doing something, but rather I make the distinction mostly for analytic purposes, so as to distinguish which kinds of illocutions are happening where within the criminalisation claim. But both illocutions – asserting and declaring – are, in the technical sense, being performed and I do not intend to propose that 'saying' something is not also the performance of an action.

and liabilities towards Φing, which might lead us to morally judge people for Φing as wrongdoers and as worthy of moral blaming, and may have given moral reasons to refrain from the conduct in question. But there were no legal duties nor liabilities with regards to Φing *qua* criminal offence up to that point, so there necessarily has to be some kind of new normative fact – creating its own new sets of reasons, duties and obligations – being established through its criminalisation. And these normative facts are, necessarily, created *by* the act of criminalising Φ – any prior normative (in this case, moral) consequences that Φing may have generated may be relevant, but are not *identical* to, the normative consequences which are entailed by criminalising Φ. It is the creation of these normative facts which allows for us to separate the legal consequences for Φing from the moral consequences for Φing (though, granted, they are not necessarily *entirely* separate but they are, at the very least, distinct from each other). And it is in this sense that I argue that an obligation not to Φ is declared – it is created as a normative fact within the context of the law.

Finally, the fourth normative fact being declared is a new responsibility with regards to Φing. By responsibility, I mean a requirement of answerability[186] – if someone Φs, they must answer with some kind of explanation or response (which, in the context of the criminal process, may be to remain silent) for the fact that they have Φed. As Duff puts it, "I am responsible for that for which I must answer, and I must answer for that which there was reason for me not to do"[187]. These reasons are precisely the ones being declared by the act of criminalising Φ – the obligation not to Φ and the liability for breaching that obligation – and it is in virtue of these reasons that

[186] Here I follow the distinction between responsibility as attributability and as answerability presented in Massimo Renzo, 'Responsibility and Answerability in the Criminal Law' in RA Duff and others (eds), *The Constitution of the Criminal Law* (2013) 209.
[187] Duff, *Answering for Crime - Responsibility and Liability in the Criminal Law* (n 1) 22.

anyone who is found Φing can be brought to answer for doing so in a criminal process. Again, this declared normative fact does not need to be explicitly stated in the speech-act, but it is also a constitutive part of it as a result of being declared in the context of the criminal law. That is, since we *already* have other rules that confer powers to legal officials to bring in people who are suspected of committing a criminal offence to answer, and this answering must be done within the criminal process, then in the act of making a conduct count as a criminal offence we are *simultaneously* declaring that anyone who is suspected of committing that offence will now have a responsibility to answer for it.

Duff's distinction between liability and responsibility is also useful here, whereby "responsibility is a necessary but not a sufficient condition of liability. I am liable to conviction or blame for X only if I am responsible for X, but I can be responsible for X without being thus liable."[188] I have mirrored it in distinguishing the liability being created from the responsibility being declared – the former only establishes what kind of consequence an offender will potentially face and the latter creates the normative facts that establishes the process by which an offender will be required to answer for their conduct (which may ultimately lead to facing conviction or not). Again, we might argue that there are also other kinds of responsibility at play, like moral responsibility for moral wrongdoing, but those are not the responsibilities that are being declared in the act of criminalisation – the answering for committing a crime being declared here is specifically letting people know that they will (or may, depending on the level of discretionary powers in the relevant jurisdiction) hereby be subjected to the criminal justice system if they are found (or suspected) to have Φed.

[188] ibid 20.

These are all normative facts which *did not exist* until the criminalisation claim was uttered. Even if we were to argue that the law is, in some important sense, only reporting purported moral obligations or that legal obligations are reducible to such moral obligations, it does not make sense to see the act of criminalising as a mere form of finger-pointing to moral obligations – it still needs to, on the basis of such obligations, create new ones that are triggered by putting the criminal law in motion when an offence occurs. And these new obligations, liabilities and responsibilities are declared – are *performed* in speech-act theory terms – by criminalising the conduct in question.

I have not yet spoken as to where wrongfulness could potentially appear in a criminalisation claim. This is because, I argue, wrongfulness of any kind (including moral) is better seen as an assertion – as a reporting of the fact that the criminalised conduct is wrong – rather than as a declaration. I will explore this possibility in the next section.

3.3 Criminalisation Claims: Asserting Wrongfulness

As I stated earlier, things are being both asserted and declared in the act of criminalising a conduct. Let us begin with what is being asserted. I propose that there is one assertion being made through a criminalisation claim – the claim asserts that the conduct is, in some important sense, *wrongful*. By the act of criminalising Φ, we are reporting as a matter of fact that Φ is wrongful. I will have more to say about what kind of wrongfulness is being asserted later, but for now it is enough to establish that there is *some* kind of wrongfulness being asserted – in most of the current literature, as we shall see, this wrongfulness is understood as of a moral kind. Again, this reporting is of a factual matter – it commits the speaker (whoever that may be in practice) to the sincere belief that Φ is wrongful, and to the truth value of that

assertion: it is true that Φing is wrongful.

A quick clarification is necessary at this point. The commitment a speaker takes on by making an assertion is regarding what the speech-act is expressing, not to *actually* holding the belief in what is being expressed. As Searle explains, when someone performs an illocutionary act with a propositional content, like an assertion, "the speaker expresses some attitude, state, etc., to that propositional content",[189] and that this holds "even if he is insincere, even if he does not have the belief, desire, intention, regret or pleasure which he expresses, he nonetheless expresses a belief, desire, intention, regret or pleasure in the performance of the speech act."[190] This becomes obvious when we make explicit how this contrast works in a sentence. If I say "it is raining outside, but I do not believe it is raining outside", the sentence seems linguistically unacceptable – something is off or infelicitous about that assertion. Of course, it is perfectly possible that, in reality, I do not *actually* believe that it is raining outside, and it is also possible that I insincerely assert that it is raining outside. The point, however, is that if I said, "it is raining outside, but I do not believe it is raining outside", that statement can no longer perform its illocutionary point of asserting a propositional content correctly. Thus, if the speech-act states "it is raining outside", that assertion commits the speaker to expressing a belief in the truth value of that statement, *even if* the speaker does not *actually* hold that belief.

Following Cornford,[191] there are two potential kinds of arguments that can be made to propose that the act of criminalising a conduct communicates some kind of portrayal of the conduct as wrongful. The first is the familiar idea

[189] Searle (n 163) 4.
[190] ibid 4.
[191] Cornford (n 15) 629–631.

from legal theory that crimes are a kind of legal wrong,[192] which I have echoed in the declared obligation discussed in the previous section, and the Razian idea that the law claims legitimate authority to make compliance with said legal obligations morally obligatory.[193] The second kind of argument relies on looking at the institutions of the criminal law, particularly the liability to punishment I discussed in the declared normative facts, and argue that those institutions only can make sense if the conduct which makes one liable to punishment is being portrayed as a wrong. As Edwards puts it, the "fact that offenders are liable to *punishment* cannot but further imply that as far as the law is concerned, offenders should not so act."[194] This same strategy is used by Husak in order to arrive at a wrongness constraint (which I will explore in detail in chapter 4), where he looks at the doctrines of the general part of the criminal law to establish moral wrongness as a constraint on criminalisation.[195] But, as Cornford points out, since there is no express mention of the wrongness of the criminalised conduct, if this wrongness *is* being conveyed in criminalisation "it must be that criminalisation has a certain symbolic *meaning*: We share the understanding that criminalisation conveys a judgement that the targeted conduct is wrongful."[196] This is, precisely, what I will argue for in this section.

One might object to this picture. One may think, as Allott and Shaer propose, that "the initial promulgation of a statute is a matter of enactment, not description."[197] That is, when a conduct is criminalised through legislation, the act by which a conduct is described as a particular offence, as

[192] HLA Hart, *The Concept of Law* (3rd edn, Oxford University Press 2012).
[193] Joseph Raz, *The Authority of Law, Essays on Law and Morality* (2nd edn, Oxford University Press 2009) ch 2.
[194] James Edwards, 'Coming Clean About the Criminal Law' (2011) 5 Criminal Law and Philosophy 315, 320. Emphasis in original.
[195] Husak, *Overcriminalization: The Limits of the Law* (n 1) ch 2.
[196] Cornford (n 15) 631.
[197] Allott and Shaer (n 175) 361.

Marmor puts it, "is not a description of how things are in the world, but rather, a prescription that one ought not to Φ in C."[198] That being the case, it may seem strange to think of the act of criminalisation to include some kind of description – even if it is an implied one – of the fact that a conduct is wrongful. There is a sense in which the 'enacting' rather than 'describing' point is trivially true – the enactment of a criminal offence creates the legal category of the offence being enacted, which was not part of the world previously (what I called the labelling fact). But could the same point be made about the possibility of including wrongfulness into a criminalisation claim? In other words, can there be a criminalisation claim that does not, at least implicitly, include some kind of assertion of the wrongfulness of the conduct being criminalised?

One way to answer this is to apply Grice's 'cancellability' test, particularly with regards to contextual cancellability whereby, as Allott and Shaer put it, "a change in the context of the utterance effaces implied content but no the encoded content."[199] The basic idea is that if we were to alter the context of the claim being made, the original implication of the utterance stops making sense. So, for an example relevant to this discussion and which I used earlier, to preclude wrongfulness from being a part of the criminalisation claim the following cancellation would need to make logical sense:

"(1) If a person ("A"), with A's penis-

(a) without another person ("B") consenting, and

(b) without any reasonable belief that B consents,

penetrates to any extent, either intending to do so or reckless as to

[198] Andrei Marmor, *The Language of Law* (Oxford University Press 2014) 64. Here, 'C' stands for 'circumstances' that provide a particular factual context for Φ.
[199] Allott and Shaer (n 108) 362, citing Paul Grice, 'Further Notes on Logic and Conversation' in Peter Cole (ed), *Pragmatics* (Academic Press 1978).

whether there is penetration, the vagina, anus or mouth of B then A commits an offence, to be known as the offence of rape.

(2) *But* this does not mean that rape is wrong."

If (2) seems counter-intuitive, then it seems strange to exclude the assertion of wrongfulness from the speech-act. It is, at the very least, an important implication given by the context in which the criminalisation claim is uttered, and taking it away from that context seems to give us an incomplete meaning for what is being said. It makes more sense to understand the utterance of the above enactment as it including the idea that raping is wrongful in some important way. And the way in which this is done, as I stated earlier, is indirectly. It is not by explicitly stating the wrongfulness of rape in the utterance of s.1, but rather it is understood from the fact that rape is being subjected to the declaration of all the normative facts I discussed in the previous section. And, importantly, it is not done by a separate illocutionary act, but rather in the original act of stating the contents of s.1 there is the indirect illocution of asserting that wrongfulness. Thus, the assertion of wrongfulness, I argue, is an *implied* assertion made as part of that indirect illocution.

How does this implied assertion work? What we need to establish is how the context in which the criminalisation claim is uttered allows for the hearer of it to rationally infer that an assertion of wrongfulness is being conveyed. The implicature here is a contextual one, so a change in the relevant aspect of the context would mean that the implicated content changes. A good way to figure this out is to look at what aspects of the context in which a criminalisation claim is uttered, if they were changed or eliminated, would negate the implicated content of the claim. As stated earlier, one possible

argument in favour of the idea that wrongfulness of the conduct being criminalised is being asserted in that act is linked to the presence of criminal punishment. The fact that liability to punishment is being established for those who Φ signals to hearers that Φ is wrongful. Again, this signalling would be done through implicature as there is no explicit statement of the wrongfulness of Φ, so if we removed criminal punishment from the context of the utterance of a criminalisation claim – and the argument were correct – then the implicated content should change. Using the same example as before, this would look something like the following:

(1) If a person ("A"), with A's penis-

(a) without another person ("B") consenting, and

(b) without any reasonable belief that B consents,

penetrates to any extent, either intending to do so or reckless as to whether there is penetration, the vagina, anus or mouth of B then A commits an offence, to be known as the offence of rape.

(2a) *But* those who commit rape shall not be liable to criminal punishment of any kind.

If my argument that wrongfulness is being implicitly asserted and can be inferred from the presence of the declared normative facts is correct, the presence of (2a) should take away any contextual information that allows us to infer, from the criminalisation of rape by s.47(1), that we are being implicitly told that rape is wrongful. And this seems like a plausible conclusion to reach – if there is no possible punishment for rape, then at least in the eyes of the law, one could infer, it is not being portrayed as wrongful. I propose that the same is true about the other declared normative facts I argued for in the previous section – declaring a legal obligation not to Φ and that there is a

responsibility to answer for Φing, gives us enough background information or context to plausibly infer that Φ is, in some way, being portrayed as wrongful. Note, however, that I am speaking here of the broad sense of wrongfulness I introduced in chapter 1, because as we shall see later, it is still an open question whether the particular wrongness being implicitly asserted is of a moral kind.

Criminalisation claims, then, are doing both assertive and declarative work – or have both kinds of illocutionary force, in Searle's terms – which are happening at the same time. It is clear, also, that if moral wrongness is to appear within the declarative stage, it must do so as the implicit assertion that the conduct being criminalised is morally wrongful. Further, there are both descriptive and normative facts being declared in a criminalisation claim. The descriptive fact is the labelling of Φ as a particular category of an offence, which may or may not be present in the criminalisation claim itself. The normative facts being declared – the obligation not to Φ, the liability for Φing and the responsibility for Φing – are new, legal normative facts that come to exist by the speech-act that criminalises Φ. The question that remains now is the following: if criminalisation claims are implicitly asserting the wrongfulness of the criminalised conduct, how do we know if that wrongfulness is of a moral kind? This is where the tools provided by metaethics will become useful.

3.4 Applying the metaethical questions to criminalisation claims

In this section, I will discuss how the three kinds of metaethical questions I posed in chapter 2 can be applied to criminalisation claims, specifically regarding the implicit assertion of wrongfulness that I argued for above. My purpose in doing so is to show how our normative theorising of what is being

said when criminalising a conduct can be improved by including these questions into our theorising, and this will become clear when we see what kinds of issues can be discussed by considering metaethical implications of using moral talk. The main question that needs to be addressed, by using the tools of metaethics, is whether it makes sense to think that the implicit assertion of wrongness that is being made by a criminalisation can be properly inferred to be of a moral kind, and how one goes about arguing for or against this view.

3.4.1 The Conceptual Question for Criminalisation Claims

Let us start with the conceptual question. Outlining what it means for the wrongfulness being implicitly asserted to be 'moral' will help determine quite a few issues around how we understand wrongfulness within the context of a criminalisation claim. As I explained in chapter 2, if the wrongfulness being conveyed is of a moral kind, this will mean that the wrongfulness is authoritatively normative. But importantly, since there are other normative domains that can also potentially be authoritative, we need to be clear as to what, metaethically speaking, gives the authoritative normativity of moral wrongness its *distinctive* authority. Once again, I will not provide a definitive answer to this question here, but I do wish to show how and why it is important to seek answers for our theorising of this stage of criminalisation.

An important point to note here is that for criminalisation claims to specifically be asserting moral wrongfulness entails that the moral character of the wrongfulness is also *implied* when making the claim. Once we accept the idea that a criminalisation claim conveys an assertion of the wrongfulness in a broad sense of the conduct being criminalised, we then need to explain how a hearer can make the rational inference that that wrongfulness is

specifically moral. Therefore, for the wrongfulness in question to also be of a moral kind, this must *also* be part of the implicature one can infer from the fact that a conduct is criminalised. Though I agree with Cornford that this poses an empirical question,[200] I argue that there is also a semantics question here to be answered: regardless of whether people *actually* have this shared understanding, is the speech-act itself conceptually able to be felicitous in conveying this implication as it stands? And to answer it, we require a prior answer to the conceptual question of what counts as morality in this context, to then determine whether that conceptual understanding of something being 'morally' wrong can properly be considered as a part of the context or background information that allows both the hearer of the criminalisation claim to infer that the wrongfulness being asserted is specifically moral, as well as the speaker to understand that when making a criminalisation claim, a natural implication of doing so is conveying the meaning that the conduct being criminalised is morally wrong.

Let us consider this issue by looking at the kind of answer that we could get by using the declared normative facts of criminalisation as our context for inference. For the sake of argument, let us accept that we *can* determine that the implicit assertion of wrongness is of a moral kind. What kind of morality, from a metaethical perspective, would make sense to attribute to that inference? In other words, if we are able to tell that a wrong is of a moral kind because the conduct has associated to it a liability to punishment, an obligation not to perform the conduct and a responsibility to answer for doing the conduct, what does the concept of morality that we can extract from those normative facts being there look like? From asking ourselves this, we can reach the following conclusion: accepting that the declared normative facts of a criminalisation claim are a good basis for inferring the moral quality

[200] Cornford (n 15) 631.

of the asserted wrongfulness entails a view that 'moral wrongness', from a metaethical perspective, can be identified by the appropriate responses to moral wrongdoing. That is, it involves accepting a version of *response-dependence* for morality,[201] because if we are able to tell something is morally wrong in view of a response to it, then that means the response needs to be assumed as a reliable indicator of that wrongness. So, if something like liability to punishment were to play that role, we would need to accept that a hearer of a criminalisation claim can see that a conduct now has a liability to punishment, and say to themselves something like "this is the right kind of response to moral wrongdoing, so in creating that liability to punishment I am probably being told that the conduct in question is morally wrong".

This is a plausible metaethical position to hold, but it is a contested position. Two arguments against it are worth mentioning here. The first comes from Blackburn, who argues that trying to specify the appropriateness of the responses in a satisfactory way faces insurmountable difficulties, particularly regarding using the disputed concept in the specification of the appropriateness for the response.[202] Similarly, the second argument comes from Johnston, who argues that there is a 'missing explanation' in response-dependent accounts of morality, because the appropriateness of the response can be traced back to the manifest properties of the object, but this explanation is 'missing' in explaining the property to the response.[203] As Gundersen explains it, referring the argument to the case of accounts of colour as response-dependent, "if something is red because it looks red, it

[201] For an overview of this line of thought, see Miller (n 19) ch 7; and for a detailed breakdown of how these metaethical stances work see Eline Busck Gundersen, 'Making Sense of Response-Dependence' (University of St Andrews 2006).
[202] Simon Blackburn, 'Circles, Finks, Smells, and Biconditionals', *Practical Tortoise Raising* (Oxford University Press 2010).
[203] Mark Johnston, 'Are Manifest Qualities Response-Dependent?': (1998) 81 Monist 3.

can't also be the case that it looks red because it is red."[204]

Finally, an added difficulty to this kind of response-dependent view has to do with how it can provide an explanation for the distinctive authority for the normativity of morality. That is, if the response to moral wrongdoing is the right marker for determining the distinctness of moral kind, then that would also mean that this is connected to the distinctness of morality's authoritative normativity. This is going to be particularly challenging, considering the fact that this kind of view will have great difficulty in surpassing Mackie's queerness argument, briefly discussed in chapter 2. If it were true that the authoritativeness of morality can be explained by the response it produces, we will need to be able to explain *where* that response comes from, in a way that surpasses the questionable nature of there being a 'special capacity' for us to respond to moral wrongdoing adequately and consistently in the same way. Since this seems like a challenging task, it would seem that the response-dependent account of moral authoritativeness "does not appear to be of the right kind to enable an analogous dispositional theory of moral value to successfully undermine the argument from queerness."[205]

Tadros has recognised the issues of response-dependent accounts in the context of criminalisation theorising, specifically regarding providing an account of wrongdoing.[206] He summarises the problem of these kinds of accounts by stating that "if the aptness of responses to morally wrongful acts depends on certain felicity conditions being fulfilled we need an independent account of these felicity conditions".[207] This independent account, in turn, is best sought in the field of metaethics, because we will

[204] Gundersen (n 201) 9.
[205] Miller (n 20) 141.
[206] Tadros, *Wrongs and Crimes* (n 1) s IV-VI in Chapter 2.
[207] ibid 23.

need to determine what kinds of facts, and what relationship between those facts and what we say about them, determines the aptness of a response. That is, if we want to propose that the distinctness of moral wrongdoing is the kind of appropriate response it elicits, we will need a metaethical explanation for how that appropriateness comes to be. So, if one is interested in proposing a theory of criminalisation as a communication of moral wrongness, it would be beneficial to be aware of the metaethical assumptions that are required to make an argument like the one I presented earlier, and be prepared to answer the critiques of said assumptions.

3.4.2 The Ontological Question for Criminalisation Claims.

The main issue the ontological question will help us elucidate is regarding what needs to be the case with regards to the propositional content of the assertion of wrongfulness of the conduct being criminalised for that assertion to be accurate. I say accurate and not felicitous, because it is clear that one can successfully make an assertion as a speech-act, while also asserting something that is false – the speech-act itself of asserting a fact still succeeds, but its propositional content is what fails in some way. In this sense, even though I accept that a criminalisation claim can felicitously assert the wrongfulness of a conduct even when the fact being asserted may be false, the point of the ontological question will be to decide what needs to be the case for that assertion to be true. There might be many reasons for why one might think an assertion of wrongfulness through criminalisation should be accurate in this sense, one of which is a moral argument in itself – it is morally wrong for the state to criminalise conduct that is not actually morally wrongful, or in other words to assert an untruth – but as Simester and Von Hirsch point out, there might be other moral reasons such as the fact that, as they put it, "people have a moral entitlement not to be designated, officially,

as miscreants when they do no wrong."[208] More generally, however, the fact that an assertion of wrongfulness might be false will depend on what we think the requirements are for that assertion to properly track whatever fact it is purporting to assert, and there may be practical, epistemic, and maybe even political reasons for why it may be important or valuable to make sure that the assertion being made is true.

So, it becomes important to determine what exactly is being asserted when a criminalisation claim asserts the wrongfulness of a conduct. Once we have committed to an answer to the conceptual question of what 'moral' wrongfulness entails, we now need to understand what is the wrongfulness itself that is being asserted through the claim – from knowing this, we can then determine what are the requirements for that assertion to obtain. When I say 'obtain', I leave open the possibility of allowing different theories of truth to determine specifically what are the truth conditions for a fact to obtain,[209] and instead what I will be focusing on is the more basic idea that, regardless of how one thinks truth conditions set the truth value of an assertion, *some* kind of reference to a particular thing in the world will be required in order to know whether what is being said is, in whichever sense one prefers, true. 'Snow is white' still needs to be talking about something, in this case snow, in order for us to assign a truth value to it, and we will need to know *something* about how we determine the existence of snow in order to talk about it – that something is what the ontological question I pose here is interested in.

Let us begin by laying some additional groundwork on how assertions work. Assertions are what philosophers of language have called constative speech-acts: they have a propositional content that has a truth-value, and that

[208] Simester and Von Hirsch (n 1) 19.
[209] For a useful introduction, see Richard Kirkham, *Theories of Truth, a Critical Introduction* (MIT 1992).

has a word-to-world direction of fit, whereby the success of the speech act requires that the content of the utterance matches a pre-existing fact.[210] As explained earlier, assertions also commit the speaker to a belief in the truth-value of the propositional content being reported – saying that *p* is the case seems to be "bound up in a special way", as Grice puts it,[211] with the fact that the speaker believes that *p* is the case. As discussed earlier in the chapter, this 'bounding up' is best explained as Searle's idea that an assertion commits the speaker to expressing a belief in *p* in the act of assertion, even if they do not actually hold the belief in *p*.

What are the success conditions for assertions that refer to objects generally? Again, Searle gives us a useful set of necessary conditions for what he calls a "fully consummated reference",[212] which is "one in which an object is identified unambiguously for the hearer, that is, where the identification is communicated to the hearer."[213] The necessary conditions he identifies for these kinds of references are:

> "1. There must exist one and only one object to which the speaker's utterance of the expression applies […], *and*
> 2. The hearer must be given sufficient means to identify the object from the speaker's utterance of the expression."[214]

So, if the assertion of a moral wrong is to be a fully consummated reference, it needs to be the case, according to Searle, that there is one and only one property of moral wrongness to which the speaker's utterance

---

[210] See Ch. 10 "Assertion and Other Speech Acts" in Zoltán Gendler Szabó and Richmond H Thomason (eds), *Philosophy of Language* (Cambridge University Press 2019).
[211] HP Grice, *Studies in the Way of Words* (Harvard University Press 1991) 42 <http://hdl.handle.net/2027/heb08428.0001.001> accessed 21 October 2022.
[212] Searle (n 164) 82.
[213] ibid.
[214] ibid.

applies, and the hearer must be given sufficient means to identify that property from the speaker's utterance. Any attempt at trying to determine whether this is the case or not, will force us to discuss the metaethics of moral wrongness, particularly regarding the ontological nature of that property. For example, a moral realist will be inclined to accept that this is the case, because they will hold that – provided one is looking for the right kind of property, depending on whether one is a naturalist or a non-naturalist – moral wrongness is only one kind of property in actions, and one can make truth-apt statements about those properties. But any other ontological position, within the spectrum of anti-realism or the forms of constructivism discussed in chapter 2, might not agree with the realist's assumptions. For example, a moral error theorist will not think that this kind of assertion is even possible, because there is nothing instantiated in the world for the assertion to refer to. An expressivist, in contrast, will think that one can make an utterance that states moral wrongness, but that the illocutionary point is *not* to be an assertion, but rather an expressive[215] – a communication of a non-cognitive mental attitude about what is being talked about. If this is the case, then the expression of moral wrongness is not trying to track anything about the world, but rather about the speaker and their conative attitudes. So, we cannot really evaluate whether there *can* be a fully consummated reference in an assertion of moral wrongness without being clear on what is the ontological nature of the property that is being referred to, if at all.

Perhaps, however, the assertion of moral wrongs might still be *successful*, rather than a fully consummated reference, where the former is, in Searle's terms, successful "in the sense that we could not accuse the speaker of having failed to refer—even if it does not identify the object unambiguously for the

[215] See Searle (n 163) 15–16.

hearer, provided only that the speaker could do so on demand."[216] This may well be the case for moral wrongs – one can assert them successfully even if ambiguously – but the key for that success is that the speaker *needs* to be able to disambiguate what they are referring to. Again, any attempt at this disambiguation will require the speaker – and, therefore, us as theorists of criminalisation – to engage in the metaethical investigations that will allow us to properly disambiguate the nature of the property that the implicit assertion of moral wrongfulness is trying to refer to.

3.4.3 The Epistemic Question for Criminalisation Claims

Finally, let us now look at the epistemic question. This question will be a very relevant issue in the next chapter, where I will discuss the deliberative stage of a criminalisation process, since if we are to propose that a deciding factor regarding the decision to criminalise a conduct has to do with moral wrongfulness, then it is vital to examine the different ways one might understand how a decision-maker gains epistemic access to that wrongfulness in order to include it in their deliberation process. Nevertheless, there are still some useful insights we can gain from applying the epistemic question to the declarative stage of criminalisation, regarding how epistemic access to moral wrongfulness might be gained *from* the speech-act which criminalises a conduct.

One potential issue that the epistemic question might help us elucidate has to do with the possibility for moral learning from criminal offences. This is because if criminal offences do indeed include an assertion of moral wrongfulness in them, then potentially the hearer of a criminalisation claim could learn about the moral character of the criminalised conduct from the

[216] Searle (n 164) 82.

fact that it has been made into an offence. Though this is a similar issue to the one I mentioned when looking at the conceptual question, on how hearers might have different conceptual understandings of what morality entails, the issue I am referring to here is slightly different. Rather than examining whether the speaker and the hearer share a common understanding of what morality is, the epistemic question I pose here questions whether a hearer can potentially gain epistemic access to the fact of the moral wrongfulness of Φ by it being criminalised. In other words, can citizens acquire moral knowledge from what is being said when a conduct is criminalised?

One way to argue that this is possible would be to point to the connection between the wrongness being asserted and the punishment that assertion leads to. This is how some theorists have approached the issue and have attempted to reconstruct, as Kleinfeld argues, the architecture for criminal theory, since "to understand wrongdoing is to learn something about the response wrongdoing calls for, which in turn implies a position as to what punishment is and what it is for."[217] And indeed, one could argue that the response given to a criminalised conduct can serve as a kind of evidence for the wrongfulness of the conduct included in an offence, or that punishment has a kind of "attitude-shaping" that resembles a form of moral learning.[218] Though as I have shown above, this has some difficulties from a metaethical perspective regarding the conceptual delimitation of moral wrongness, let us assume for now that it is plausible to use the declared normative facts as the context necessary for inferring moral wrongness.

As I mentioned, when a criminalisation claim is made there are three

[217] Joshua Kleinfeld, 'Embodied Ethical Life and the Criminal Law' in Chad Flanders and Zachary Hoskins (eds), *The New Philosophy of Criminal Law* (Rowman & Littlefield International 2016) 40.
[218] For an older version of this position, see Johannes Andenaes, 'The Moral or Educative Influence of Criminal Law' (1971) 27 Journal of Social Issues 17.

normative facts being declared: an obligation not to Φ, a liability to particular consequences for Φing and a responsibility to answer for ones Φing. The presence of these declared facts can be argued to act as the *context* from which a potential hearer can infer the implicit assertion of wrongfulness for the conduct being criminalised. In other words, if we accept the idea that the appropriate response to wrongdoing is to make the kind of declarations done through a criminalisation claim, then we can accept that the presence of these normative facts entails an implicit assertion of wrongfulness.

There are two potential difficulties with this argument. First, if it were the case that these declared normative facts are sufficient context, then that would mean that *any* conduct which lead to the declaration of those facts would be reasonably inferred to be a moral wrong, and this does not seem plausible – there are plenty of conduct that are criminalised, where those declared normative facts are established, which we would probably hesitate to call moral wrongs in the first place. Further, following Cornford's point, there is also the very real possibility that in practice, people end up being submitted to the effects of those declared normative facts when they actually haven't acted in a way that would plausibly be described as wrong, and instead have either been put in a position where they need to plead guilty or they find themselves in a situation where there is no room for reasonable doubt from the evidence, even when a doubt that is unreasonable, form the fact-finder's perspective, is in fact true.[219] In those cases, if the argument of moral learning from the declared normative facts were true, we would still be in an epistemic position with sufficient evidence available to justify forming the belief that the defendant has committed a wrong, yet this seems clearly false. The declared normative facts in a criminalisation claim may be a useful *prima facie* indicator of the wrongfulness of the conduct involved, but they

[219] Cornford (n 15) 620–621.

are by no means sufficient evidence for it.

Second, we might run into issues regarding the justification[220] of declaring the normative facts of a criminalisation claim. The characteristic of declaratives as speech-acts is that they have a double direction of fit – they need a prior fact to exist, so as to match the content of the utterance with what justifies making the declaration in the first place.[221] That being the case, we run into a challenge if we try to explain an implicit assertion of wrongfulness through the declared facts, because those declared facts *need* there to be certain pre-existing facts *in order to come to be*. The whole reason why those normative facts are declared in the first place is that they are justified by there being a legitimate authority that enables the speaker to make these declarations.

One way to explain this is to follow a view such as that of Raz, where the prior fact that the declarative facts need to come to be is just the legitimate authority the law claims to have to make compliance mandatory,[222] and as Cornford puts it, "since crimes are legal wrongs, they are also putative moral wrongs: In the eyes of the law, it's wrong for citizens to engage in the conduct that they target."[223] In this sense, the declared normative facts are justified by the fact that the *source* of the speech-act is legitimately authorised to do so.[224]

However, this is usually not the way that theorists speak of the justification for the declared normative facts of criminalisation, particularly when it comes

[220] Here, I use 'justification' in a broad normative sense. It can be read as moral justification in some contexts, and I will signal this by explicitly stating that I am speaking of moral justification.
[221] See Searle (n 163) 16–20.
[222] See Ch.2 in Raz (n 193).
[223] Cornford (n 15) 630.
[224] Raz's particular and well-known arguments for the justification of authority can be found in Joseph Raz, *The Morality of Freedom* (Oxford University Press 1986) ch 3.

to liability to criminal punishment. I will explore these views in more detail in chapter 4, but I have already pointed out earlier some of the views that state that moral wrongness ends up acting as a justification for the creation of liability to punishment.[225] That is, subjecting a person to the liability to being punished is justified by the fact that said person has committed a morally wrongful act. However, is this *also* the case for *declaring* that liability to punishment? In other words, is the declaration of the normative facts included in a criminalisation claim justified by the authority by which the speaker utters the declaration, or is it justified by the fact that the conduct being criminalised is morally wrongful?

One answer is to say that it is both. That is, the illocutionary act of declaring normative facts is justified *as a speech-act* by the authority that the speaker holds to make the declaration, *and* the declared normative fact of creating a liability to punishment is justified *as a normative fact* by the wrongness of the conduct being criminalised. If this is true, however, the challenge now becomes determining which of these senses of justification a hearer can reasonably infer from the criminalisation claim and its declared normative facts. This is because the 'appropriateness' of the declared normative facts can be seen as both of these senses of justification: liability to punishment, for example, is 'appropriate' because the state is legitimately authorised to declare this kind of liability, *or* liability to punishment is 'appropriate' because it is the right kind of response to moral wrongdoing. Which of these senses is a hearer supposed to infer? Again, we might answer that they can infer both, but this leaves us in a strange theoretical position: the inference that a hearer makes from a criminalisation claim is, at the same time, an inference of legal wrongdoing and of pre-legal wrongdoing, from the exact same pieces of contextual information. In other words, a hearer can infer that the conduct is

[225] See the discussion of the liability fact in Sec. 3.2 above.

'morally wrong' *and* 'morally wrong according to the law', but since they can do so from the same pieces of context, these two inferences become almost indistinguishable from each other. Therefore, this leaves us with a conundrum regarding the epistemic justification of the inference that a hearer can make regarding the moral wrongness of a criminalised conduct. It seems that a hearer can make some kind of inference on the wrongness of the conduct based on the declared normative facts included in a criminalisation claim, but what exactly is providing this epistemic justification? Is it the fact that those declarations are made by a legitimate authority, like the legislature? Or is it the fact that the declarations are the appropriate markers for identifying a moral wrong? I cannot explore this issue further here, but suffice to say that getting to this analysis was possible thanks to the use of the metaethical tools I have presented previously, which again gives us a good indication that doing so can help us improve our normative theorising of criminalisation.

In conclusion, in this chapter I have proposed a model for understanding the act of criminalisation as a speech-act, and to identify within that model where moral wrongness could be argued to appear. I have shown that the best way to theoretically place moral wrongness in what is being said through criminalisation is as an implicit assertion, that can be inferred by hearers of the speech-act by looking at the declared normative facts I identified – obligation not to do the conduct, liability to punishment, and a responsibility to answer – as a portrayal of the criminalised conduct as wrongful. I then analysed some of the metaethical implications of holding this position, both at a conceptual, ontological and epistemic level. The purpose of doing this has been to show that, though it is a plausible position to hold, arguing that moral wrongness is part of what is being said through criminalisation needs to be theorised with metaethical investigations in mind. Otherwise, the foundations on which our understanding of criminalisation as a form of

communication stand on are, at best, incomplete.

CHAPTER 4 – THE DELIBERATIVE STAGE OF CRIMINALISATION

If we accept one of the main assumptions upon which my proposed declarative stage rests — namely, that we choose to criminalise conduct through a particular act that makes it the case that a particular conduct is considered a criminal offence (i.e. we make a criminalisation claim) — then it makes sense to then ask ourselves how we should decide to perform that act of criminalisation. In other words, if the act of criminalising is indeed a choice — and I will assume it is — then there must be a process by which we choose to do that act. This process is what I call the *deliberative stage* of criminalisation: it is the process by which one decides that a conduct should be criminalised. It is at this stage of the criminalisation process where a potential formal decision-maker (usually a legislator, but not necessarily) will come up with different reasons to justify the criminalisation of a conduct. It is also the stage of the process which has received much scholarly attention from criminalisation theorists, where great efforts have been made to construct normative theories of criminalisation which propose principles that should guide these decisions about what conduct to criminalise.

The chapter begins by explaining what the focus of the analysis is within the deliberative stage, and shows the different ways in which moral wrongfulness has been understood when used in criminalisation theory. It then looks at five recent works on criminalisation theory that explore the role of moral wrongfulness: Michael Moore's "Placing Blame", Antony Duff's "The Realm of Criminal Law", Doug Husak's "Overcriminalization", Andrew Simester and Andreas Von Hirch's "Crimes, Harms and Wrongs", and Victor Tadros's "Wrongs and Crimes". Each one is explained, with particular emphasis on how each theorist proposes both how a decision on

criminalisation should be made, and how moral wrongfulness should play a role (if any) in that decision-making process. For each theorist, I present a reconstruction of what a deliberation process for criminalisation should look like in their view. By reconstruction, I mean that I will be piecing together what a criminalisation deliberation process looks like for each theorist, which sometimes will be a reproduction of what the theorist has explicitly stated, and other times will be a construction of a deliberation process based on how they discuss criminalisation decision-making, even if they might not provide an explicit principled structure for the deliberation process. Then, I apply the metaethical questions posed in chapter 2 to each theorist's model of deliberation, and reconstruct potential answers that can be inferred from how each theoretical deliberation process is presented. The chapter concludes with a discussion of why, as we will come to see, the different answers that can be reconstructed for the metaethical questions matter to the proposed deliberation processes for criminalisation. I argue that different answers to the metaethical questions lead to different epistemic methods, and therefore different epistemic justification requirements, for the decision to criminalise conduct. This is particularly salient regarding the epistemic process that a decision-maker is supposed to engage with when trying to establish the moral wrongness of a putative criminalizable conduct, such as what kind of evidence will the decision-maker be looking for regarding that wrongness, and what facts will they be looking for to determine the relevance of certain properties of the conduct in question.

4.1. Moral Wrongfulness and the Deliberative Process

My focus in this chapter will be on those proposed principles which contribute to deciding about criminalisation that use moral wrongfulness, rather than on the deliberation process itself. Again, the point of the

deliberative stage is not to describe how people make a decision about criminalisation, but to look at how one would go about establishing a normative framework under which that decision should be made. It is, precisely, what many criminalisation theorists are trying to offer as part of their analysis of how criminalisation should work. In this sense, the point of this stage is conditional: if one is trying to propose a normative framework for deliberation as to whether a conduct should be criminalised or not, then this is another stage at which the concept of wrongfulness can appear. It will be a separate question whether, in practice, an actual decision-maker about criminalisation utilises the theoretical framework that is being proposed, and that is not the question which I intend to engage with here. Instead, my purpose is to look at the theoretical framework itself, as a proposed deliberation framework, and how wrongfulness can be used as a part of that framework. The contribution I am making with this chapter is to demonstrate the effects of metaethical commitments on that decision-making process, and the point of doing so is also not a descriptive claim of how a particular decision-maker empirically thinks, but rather on how these different metaethical frameworks shift how the deliberation process would have to work.

However, there are a couple of things that I wish to address about how I have defined what descriptively is happening at this stage, specifically with regards to the 'one' and the 'should' that is involved in the decision. Regarding who the decision-maker is, my claim is that the proposed justifying reasons for or against criminalisation are being proposed for whoever may in practice fill the role of decision-maker – it may be that in practice, the role will be filled by the legislature, a judge, a government agency or any other relevant legal official that has the relevant authority to criminalise, but the point is that *anyone* in that role would be acting as the decision-maker to

whom putative justifying reasons to criminalise are relevant. It seems that most criminalisation theories of interest to this project are discussing principles that ought to guide decisions about what to criminalise in an ideal sense, and are not too focused on the particular actor that makes the decision, or they are content with assuming that their principles are addressed, at least most of the time, to the legislature.[226] Note, however, that the relevant decision is whether to define a conduct as a criminal offence, which in the context of my model is the formal performance of a criminalisation claim. There are important empirical and social questions to consider (as well as further potential normative questions) regarding the characteristics of the people occupying roles that lead to substantive criminalisation decisions *once the conduct has been defined* as a crime, like police officers and prosecutors, but those are not the kinds of decisions that are the target of the deliberative stage I propose.

As for the 'should' involved in the decision, the sense I refer to for when a conduct 'should' be criminalised in this context is not from the perspective of some particular normative domain — it is not, for example, a moral should — but instead is a more modest 'should' which results from the deliberation process. In other words, the only way in which we 'should' criminalise a conduct, in the sense I will use here, is in view of the considerations we have come up with in our deliberation process and, once we have considered them all (however that may be done), we come to a particular conclusion that we should (or should not) criminalise a conduct. That is to say, there are many ways in which 'should' can be understood as a normative concept, among which is a moral understanding, but my interest is in the deliberative process that would be required however we choose to understand the sense of

[226] See, for example, Simester and Von Hirsch (n 1) 3–6.; Moore, *Placing Blame - A General Theory of the Criminal Law* (n 1)., ch. 16 as addressed to "legislative aim". For a criticism of this assumption, see Duff, *The Realm of Criminal Law* (n 1) 39–43.

'should' being employed in the conclusion. Therefore, I am not interested in looking at the moral judgment one can make of the deliberation done in the context of deciding whether to criminalise or of the conclusion reached by it – as in asking whether it is morally right or wrong to criminalise a conduct because it is morally wrongful. Instead, I am interested in looking at how the deliberation is shaped by a requirement to consider the moral wrongfulness of the conduct in question.

As already stated, the focus for this analysis will be on theories that use moral wrongfulness specifically. This can be done either by offering moral wrongfulness as a positive reason in favour of criminalising a conduct, or by using moral wrongfulness as a negative constraints o criminalisation. Duff has called these views Positive and Negative Legal Moralism[227] respectively and I will follow this denomination here too.

The tradition of using moral wrongfulness as a relevant reason for or against criminalisation is long,[228] and it is easy to see why that might be seen by some theorists as being the case – as Duff has put it, "it is an obviously relevant objection to a proposed criminal statute that the conduct it criminalises is not wrongful" and therefore the relevance of moral wrongfulness is "implicitly assumed" in the debate.[229] The main way of using moral wrongfulness is by looking at what happens in criminalisation and making a normative claim about how moral wrongfulness is relevant to what is happening when a conduct is criminalised. This, in turn, can have two different targets within criminalisation as a process, for which Tadros makes

[227] See RA Duff, 'Towards a Modest Legal Moralism' (2012) 8 Criminal Law and Philosophy 217, 218–222.
[228] For a summary of the debate around criminalisation, see Lacey (n 11) 937–941. See also Lindsay Farmer, 'Criminal Law as an Institution' in RA Duff and others (eds), *Criminalization - The Political Morality of the Criminal Law* (Oxford University Press 2014) 82–89.
[229] Duff, *Answering for Crime - Responsibility and Liability in the Criminal Law* (n 1) 81.

a useful classification when talking about "restrictive principles" for criminalisation. He identifies content-focused principles as those which claim "that only conduct of a certain kind – harmful conduct, for example – is permissibly criminalised", whereas effect-focused principles "claim that it is only permissible to criminalise some conduct if criminalising that conduct has certain effects."[230] Edwards uses similar categories to illustrate the distinction I am referring to with regards to different kinds of legal moralisms, which he calls act-centred and instrumental legal moralism (ALM and ILM respectively), where ALM endorses the idea that "the fact that Φing is morally wrong is a reason to criminalise Φing", whereas ILM endorses the idea that "the fact that criminalising Φing will probably prevent moral wrongdoing is a reason to criminalise Φing."[231] Therefore, the targets of these reasons are different: the first (content-focused and ALM) focus specifically on something *about* Φ or Φing itself, whereas the second (effect-focused and ILM) focus on something that happens *because of criminalising* Φ.

There are two ways of understanding these classifications. One is to say that what is relevantly morally wrongful for content-focused or ALM principles will be Φ itself, whereas for effect-focused or ILM principles it will be Φ as a result of another conduct, say ∏, that leads in some way, causally or otherwise, to Φ. Although ∏ might not in itself be wrongful, the avoidance of the wrongfulness of Φ is what does the work in the deliberation on whether to criminalise ∏ – the question is not 'is ∏ morally wrongful' but rather 'does ∏ lead to a morally wrongful Φ and, if so, should we criminalise ∏'. So, for example, we could use this kind of principle to explain why something like possession of uranium without authorisation should be criminalised, by

[230] Tadros, *Wrongs and Crimes* (n 1) 92.
[231] James Edwards, 'An Instrumental Legal Moralism' (2018) 3 Oxford Studies in Philosophy of Law: Volume 3 157. See also James Edwards, 'Harm Principles' (2015) 20 Legal Theory 253., where this kind of distinction got its starting point regarding the harm principle.

arguing that though possessing uranium in itself may not be at all wrongful, doing so can plausibly lead to building a nuclear explosive device that uses uranium and, in order to prevent that potentially wrongful act, we criminalise the possession of uranium without authorisation.

This is different from another kind of effect that might be of concern to theorists, which has to do with the wrongness of criminalising in the first place. One may have a certain view on what is permissible for the state to do to its citizens, which does not include the possibility of a state subjecting non-wrongdoers to punishment. It seems plausible to describe subjecting said people to punishment as an 'effect' of criminalisation that we may want the state to avoid. And one way to avoid punishing non-wrongdoers is to make sure that the conduct the state is criminalising is wrongful in the first place. Nonetheless, notice that this effect is still going to end up being linked to the normative status of Φ, in the sense that the relevant moral wrongfulness for the criminalisation principle to work will still be the moral wrongfulness of Φing. That wrongfulness, however, is being used in this case instrumentally as the basis for avoiding *further* moral wrongdoing, this time by the state. In this sense, this kind of principle requires us to accept two different but linked claims: that it is morally wrongful to subject citizens to the consequences of criminalisation if they have not committed any wrongs themselves, and that the conduct being criminalised is morally wrongful itself. Content-focused principles, it is important to note, can also share this concern with the effects of criminalisation. As Tadros puts it, "the concerns of those who defend content-focused restrictive principles are better understood as concerns about liability",[232] where liability is specifically regarding potential punishment. In this sense, the content-focused theorist wants to avoid wrongdoing by the state punishing a non-wrongful conduct, thereby making

[232] Tadros, *Wrongs and Crimes* (n 1) 92.

the wrongful character of Φ relevant to the decision-making.

It is useful to remember at this point as well that a criminalisation theory does not need to necessarily present one single 'master' reason or principle for criminalising conduct – a theory can perfectly well combine a series of different supporting reasons for (and constraints on) criminalisation, that can come from different normative domains and can be both conduct-centred and effect-centred.[233] My interest here, however, is to specifically look at theories which propose as a reason (at least one important reason) moral wrongfulness of either the conduct in question or the effects of criminalising a conduct. With all this in mind, I shall now take a closer look at some specific theories that use moral wrongfulness as an important consideration for the deliberative process of criminalisation.

4.2. Criminalisation theories that use moral wrongfulness for the decision-making process

Now, I will look more closely at some criminalisation theories that propose using the concept of moral wrongfulness to both justify and limit, in some way, the criminalisation of conduct. It would be impossible to carry out a complete survey of every contribution to criminalisation literature; the theorists I have selected are both recent and influential contributions to the theory of criminalisation. There are other historically relevant works in the field,[234] but more recent criminalisation literature has made both more prominent and more explicit use of moral wrongfulness as a criminalisation

[233] On why it might not be possible to present a single, 'master' principle of criminalisation, see the introduction in RA Duff and others (eds), *Criminalization - The Political Morality of the Criminal Law* (Oxford University Press 2014).
[234] For some classic examples, see Feinberg (n 69); J Schonsheck, *On Criminalization* (Kluwer Law International 1994); Devlin (n 47); HLA Hart, *Law, Liberty and Morality* (Oxford University Press 1963).

principle, making them better examples for analysis of the effects of metaethics in criminalisation theory.

The goal of this analysis is to understand the deliberative structure being proposed by these theorists, in order to then be able to evaluate how that structure can be affected by shifting metaethical commitments. I am also particularly interested in looking at how not being clear in the metaethics being assumed in these theoretical structures can lead to misunderstandings or to different representations of how these theories would work. This is to show that not only can we have different normative debates based on how we structure the deliberative process, but we can also have different meta-normative understandings of how the structure works, and those understandings are worthy of debate.

4.2.1. Michael Moore – Placing Blame

Michael Moore proposes a strong version of legal moralism, which is embodied in his book Placing Blame. In it, Moore develops a conception of criminalisation that is strongly based on moral wrongfulness as a requirement of the function of the criminal law. He argues that the only function of the criminal law — the only "intrinsic good" it pursues — is to achieve retributive justice.[235] He believes this in view of two considerations. First, he mentions the fact that there is a tension between crime-prevention and retributive goals, because retributive justice demands that punishment be inflicted purely because the offender did the offence. Any reason other than 'they deserve it' does not achieve retributive justice. Second, in view of the structure of "Anglo-American criminal law", the criminal law can better

[235] Moore, *Placing Blame - A General Theory of the Criminal Law* (n 1) 23.

achieve retributive justice rather than crime prevention.[236] Therefore, if the only function of criminal law is retributive justice – and Moore believes this is so – then those who *deserve* punishment must get it. In order to deserve punishment, two conditions are required: a wrongful action and that said action is done culpably. So, two doctrines are required, Moore argues: one that specifies acts as wrong, and one that both establishes the conditions under which an act is done wrongfully and the conditions for establishing culpability for doing so.[237]

From this framework, Moore then goes on to construct what he calls a 'theory of proper legislative motive', which establishes the legitimate aims (i.e. which states of affairs) a legislator can achieve by their legislation.[238] His proposal is a strong 'perfectionist' theory based on the moral wrongfulness of an action, whereby "because an action is morally wrong is always a legitimate reason to prohibit it with criminal legislation."[239] Even if there may be space to believe that the moral nature of the action might be contested, Moore states that a "legislator should never restrain himself/herself from following his/her own best theories of what is morally wrong just because they are his/her own theories or just because they are theories of what it is morally wrong to do."[240] He allows for the possibility that other goods may outweigh the good that is achieved by the prohibition of immorality, but immorality is always a valid reason in favour of prohibition.[241]

In this sense, when it comes to potential reasons a criminalisation decision-maker might include into their deliberations, Moore believes that the moral

[236] ibid 28.
[237] ibid 33.
[238] ibid 68.
[239] ibid 70.
[240] ibid.
[241] ibid.

wrongfulness of the conduct in question is what ends up doing all the justificatory work for the decision. As he puts it, "criminal laws are justified because and only insofar as they prohibit moral wrongdoing."[242] One way he explains the plausibility of this is by appealing to the moral intuition that most people feel towards things like cruelty to animals and the mutilating of dead bodies. Moore claims that most people 'feel' like those things ought to be criminalised, and that this must be because we think such behaviours are wrong and that society ought to legislate against those wrongs.[243]

So, for Moore the relevance of moral wrongdoing depends on which theory of punishment one follows. For him, punishment pursues the only function which is actually served by the criminal law – achieving justice through retribution. The desert that triggers retributive punishment, for Moore, is the culpable wrongdoing done by persons, and both culpability and wrongdoing are based on moral norms of obligation that tell us what to do. Wrongdoing is the breach of obligations in the actual world, whereas culpability is the breach of obligations in possible worlds created by the content of the actor's mental states.[244] This, Moore claims, is the almost complete picture to justify a legal moralist theory of criminalisation, with the only piece missing being a principle of legality which "requires that there be clear, public, consistent, prospective, general, legal rules prohibiting conduct before that conduct may be punished."[245] In virtue of this principle, then, it is the legislator who must bring moral wrongdoing into the purview of the law in order to achieve "the good of retributive justice".[246]

Curiously, Moore is one of the few legal moralists who is relatively explicit

[242] ibid 646.
[243] ibid 646–647.
[244] ibid 660.
[245] ibid.
[246] ibid 661.

about his metaethical commitments – he is happy to accept that his legal moralist position assumes moral realism, where the legal moralist legislator has "built in some version of a non-relativistic meta-ethics", on which there are right answers to moral questions that do not depend on what people think.[247] With regards to the possibility of doubts around the moral status of the conduct in question, Moore himself admits that "Usually the meta-ethical realism adopted by most legal moralists is thought to preclude the kind of doubt that gets this sort of argument going."[248] He then goes on to state, however, that "moral realists do not have to be epistemic simpletons. They do not have to believe that they *know* answers to questions like that involving the morality of incest, just because they do believe there to *be* one, perhaps general, answer. Even when moral realists have reached a conclusion in which they have some confidence, they should realize that the very possibility that they are right strictly requires the possibility that they could be wrong in these views."[249]

Moore is critical of what he calls the "absolutist view of morality", which he views as the 'most popular' version of agent-relative moral views, whereby "moral norms that make up morality are (1) short and exceptionless injunctions, such as 'thou shalt not kill'; (2) 'absolute' in the sense that they cannot be violated, whatever the consequences may be of not violating them on some occasion; and (3) applicable to what we indirectly cause as well as what we directly do through our actions, applicable to what we allow to happen as well as to what we make happen."[250] He believes that such absolutes are not contained in morality, and instead argues for a 'complex agent-relative' view of morality where consequences, under some

[247] ibid 645.
[248] ibid 662–663.
[249] ibid. Emphasis in the original.
[250] ibid 687–688.

circumstances, are relevant and must be calculated when choosing a course of action.[251]

Let us take stock of Moore's position, particularly regarding what a deliberation process is going to look like for his criminalisation theory. A potential decision-maker, according to Moore, is going to require a specifically moral deliberation process, by which the decision-maker will be applying a doctrine that requires a potentially criminalizable act to be morally wrong. In doing so, the decision-maker is looking for a positive reason that will weigh in favour of criminalising that act, so the deliberation process will very much have to focus on establishing the property of moral wrongness of the act in question, as it is, in Moore's terms, always a valid reason in favour of prohibiting that act.

Several questions are raised just by these first steps in the deliberation process. First, what does the doctrine that specifies wrongness look like to Moore? He believes that it manifests in the special part of the criminal law,[252] which "consists of those legal standards that prohibit certain behaviour" and that "one may think of […] as the criminal law's Ten Commandments".[253] Because of this, Moore argues that theorising the special part of the criminal law "must be normative, not descriptive, urging what the law ought to be, not what it is" and that "such a theory should be targeted, in the first instance, at least, at legislators, not judges, telling legislators what sorts of things may legitimately be prohibited by them."[254] Hence why Moore speaks later of proper legislative motive regarding how a legislative decision-maker ought to decide what to criminalise. Therefore, for Moore, the doctrines of the

[251] See ibid 688 onwards.
[252] ibid 33.
[253] ibid 64.
[254] ibid 67.

special part – and therefore the doctrines that a formal criminalisation decision-maker ought to follow – will need to mirror or accurately capture moral facts about potentially criminalizable conduct, for the decision-maker to correctly judge whether there is a weighty positive reason in favour of criminalising the conduct in question.

Second, what does the judgment of the existence of moral wrongness of a conduct look like to Moore? Again, since he is a robust moral realist, it seems reasonable to assume that the judgment would probably be one of ascertaining a mind-independent property of the conduct, though as we shall see below it is specifically a kind of natural property that can be found similarly to other scientific facts as an inference to the best explanation, specifically as supervening moral qualities over nonmoral qualities.[255]

So, Moore would probably expect the deliberation process for criminalisation to look as follows: the decision-maker (for Moore, a legislator) must judge whether an act contains within it a supervening property of moral wrongness. If they find this supervenience, then this acts as a weighty positive reason in favour of criminalising the conduct, as it allows for the criminal law to fulfil its ultimate normative function of achieving retributive justice.

4.2.1.a. Applying the questions to Moore

The central metaethically relevant issue with Moore's theory concerns the deliberation by which we are supposed to make criminalisation decisions. Since Moore requires a doctrine that specifies acts as wrong – and since he proposes this in the context of the criminal law fulfilling a goal of retributive

[255] See Michael S Moore, 'Moral Reality Revisited' (1992) 90 Michigan Law Review 2424, 2511 onwards.

justice – we now require an explanation of what that doctrine is supposed to look like. Moore is trying to cast the criminalisation deliberation process as an instance of moral deliberation, and hence it becomes necessary to establish a method for legal decision-makers to partake in that moral deliberation. We can combine this idea with Moore's understanding of naturalistic supervenience of moral properties and reach the following view: the doctrine which specifies an act as wrong, in the relevant sense for criminalisation, will allow for a decision-maker to correctly identify the supervening natural facts which act as token instances of moral properties.[256] With this, we can now have an idea of the kind of answer that Moore would give to the epistemic question: a criminalisation decision-maker can find the moral wrongness of a conduct by putting themselves in an epistemic position which allows them to perceive the natural facts which are supervened by the relevant moral property of moral wrongness.

Moore is, perhaps unsurprisingly, what I have dubbed earlier a robust moral realist. In particular, he defends a particularly strong ontological stance on how mind-independence works. He states:

"The ontological theses of the realist are two: first, an existential thesis that asserts the existence of the entities in question, be they numbers, legal rights, mental states, or moral qualities; second, an independence thesis that asserts the mind-independence of the entities in question. This second thesis asserts that, for example, electrons would exist even if we had no thoughts about them and, indeed, if we did not exist."[257]

In this sense, Moore's answer to the ontological question is quite clear-cut:

[256] See ibid 2516 on Moore's understanding of token identity between moral and natural property instances.
[257] ibid 2433.

moral wrongness is an objective, mind-independent property of actions. However, he has a particularly unique understanding of what, to him, counts as moral realism. As briefly explained above, Moore's realism is of a naturalistic kind, and particularly is a naturalism explained by the metaphysical claim of supervenience. Let us set aside for the moment the contested nature of the metaphysical claim of supervenience as the foundational explanation for metaphysical properties, and focus instead on the more basic claim linked to moral realism more generally. Moore's naturalism entails that the property of moral wrongfulness is accessible by the criminalisation decision-maker through the same basic *methodologies* that someone would apply when searching, for example, for electrons, neuron firings or something of the like. This does not entail, however, that those properties are *reducible* to the electrons or brain firings that the decision-maker might find in their deliberation.[258] Rather, the idea is that in finding the particular natural property that supposedly embodies moral wrongness in some way, the decision-maker is epistemically justified in explaining the presence of that property as the conduct being morally wrong (regardless of the metaphysical nature of that explanation, such as supervenience, grounding or something else).

Suppose, however, that our decision-maker is a strong anti-realist – they believe that there is nothing that exists in the world to instantiate the property of moral wrongness. It seems reasonable to believe that they would strongly disagree with Moore's theory, but on different levels of disagreement. On one level, the strong anti-realist can disagree with the whole enterprise of searching for a property 'moral wrongness', as there will be nothing to find. On another level, however, the strong anti-realist can also disagree with the

[258] For a detailed account of the differences between reductionist and non-reductionist naturalist accounts, see Miller (n 20) chs 8 and 9.

kind of epistemic method that they need to apply to find the relevant property for a decision on criminalisation. This can be regarding the naturalist claim of Moore's realism of trying to find a natural property of the conduct, or it can be regarding the seeking of a property of the conduct type itself. Instead, the strong anti-realist might suggest, they need to look for evidence of wrongness in a different normative domain than morality, or they might think that, in fact, what we need to inquire about is evidence of the evolutionary traits of humans that support the fiction of something being morally wrong.[259] Regardless of the specific suggestion, however, the strong anti-realist is not forced to abandon a theory like Moore's because it is based on a robust realist metaethics, but they will firmly disagree with Moore's picture of the deliberation process needed to figure out the wrongness of the conduct being criminalised.

4.2.2. Antony Duff – The Realm of Criminal Law

One of the most prominent and important contributions to contemporary criminalisation theory has been the work of Antony Duff. The central claim Duff makes about criminalisation — that the appropriate focus of the criminal law and criminalisation is only 'public wrongs' — stems from a robust body of work that builds a rich and comprehensive understanding of the criminal law as a whole, its aims and its role in the context of a political community. Within that body of work, his monograph "The Realm of the Criminal Law" is his most sophisticated and developed version of his criminalisation theory,[260] and will

[259] For a view along these lines, see Joyce (n 59).
[260] For some examples of his previous work, see Duff, *Answering for Crime - Responsibility and Liability in the Criminal Law* (n 1); Duff, 'Towards a Modest Legal Moralism' (n 227); SE Marshall and RA Duff, 'Criminalization and Sharing Wrongs' in Paul H Robinson, Stephen P Garvey and Kimberly Kessler Ferzan (eds), *Criminal Law Conversations* (2009); RA Duff, *Punishment, Communication and Community* (Oxford University Press 2001).

be the main focus of this section in order to determine how Duff proposes to use moral wrongfulness as a consideration or reason to be taken into account when deliberating about criminalisation.

Duff summarises his proposed principle of criminalisation as follows:

> "A. We have reason to criminalise a type of conduct if, and only if, it constitutes a public wrong.
> B. A type of conduct constitutes a public wrong if, and only if, it violates the polity's civil order."[261]

From these principles, we can notice several things. First, Duff is proposing a kind of positive legal moralism, where the fact that a type of conduct is morally wrong acts as a positive reason in favour of criminalising it. However, Duff's version adds the qualifier of 'public' to the deliberation, which entails that there is an important difference between a more general sense of 'moral wrongs' and a more narrow sense of 'public moral wrongs'. Second, Duff delimits what counts as a public wrong as a wrong which 'violates the polity's civil order', and goes on to define that civil order formally as "that institutional, legal, and social ordering through which it [the polity] defines, actualises, and sustains its guiding goals and values."[262] It is "structured by the set of goals and values through which a polity constitutes itself."[263] Hence, the moral wrong in question needs to be of the 'public' kind in order to be the business of the criminal law, and we can know if this is so by asking whether the wrongfulness in question violates the polity's civil order. To better understand these ideas, we need to see how Duff defines what civil order is, what its relationship with morality is and what role wrongs play in it.

---

[261] Duff, *The Realm of Criminal Law* (n 1) 277.
[262] ibid 185.
[263] ibid 7.

Importantly, Duff's legal moralism (as does Moore's as seen above) focuses on what he calls, following Hart, 'critical' rather than 'positive' morality, where the focus or ground of criminalisation should be actual rather than mere perceived wrongdoing.[264] Under this idea, the focus of any potential decision-maker must be not whether a conduct is regarded in fact as a moral wrong; rather, they "will need to argue that it is *rightly* so regarded."[265] In this sense, Duff also supports a form of negative legal moralism, because "we behave unjustly, and impermissibly, if we define and treat as criminal conduct that is not, or that we do not take to be, wrongful."[266] This is especially important considering the role that criminalisation as an action has, regarding specifically what is being said when a conduct is criminalised. As he puts it, "if to criminalise conduct is to portray it as wrong, this is to declare that those who engage in such conduct merit censure as wrongdoers: if the conduct is not wrongful, to criminalise it is to make a morally mistaken declaration; if the criminaliser does not even believe it to be wrongful, to criminalise it is to lie."[267]

However, Duff makes a caveat on what kind of morality he is interested in – and the kind that the criminal law should be interested in – by agreeing with theorists like Thorburn[268] and Chiao[269] that "we must theorise criminal law as part of the institutional, legal structure of the state – a theorisation that

[264] ibid 54.
[265] ibid 55.
[266] ibid 57.
[267] ibid.
[268] See Malcolm Thorburn, 'Criminal Law as Public Law' in RA Duff and Stuart P Green (eds), *Philosophical Foundations of Criminal Law* (Oxford University Press 2011).. See also Malcolm Thorburn, 'Constitutionalism and the Limits of the Criminal Law' in RA Duff and others (eds), *The Structures of the Criminal Law* (Oxford University Press 2012).
[269] See Vincent Chiao, *Criminal Law in the Age of the Administrative State* (Oxford University Press 2019). See also Vincent Chiao, 'What Is the Criminal Law For?' (2016) 35 Law and Philosophy 137.

belongs to political philosophy rather than to moral philosophy, if we understand 'moral philosophy' as concerned with 'inter-personal morality' rather than with political structures."[270] This is why he places so much focus on constructing the idea of a civil order, because it is within that context that the relevant kind of morality for the criminal law is developed, where something like a "public legal moralism" can flourish and where the legal moralist theorist "can show not merely that it is good that moral wrongdoers be punished, or called to formal public account, or that this conduces to the moral betterment of the world; but that such a practice contributes to the proper aims of a decent polity."[271]

This is why Duff makes great efforts to try and establish what counts as a polity's civil order. To start, he recognises that although the law and legal frameworks play an important role in constituting civil order, "a polity's civil order includes a social order—a set of social practices and expectations that, although they exist and are sustained within the framework of law, are not themselves legal practices, and are not directly regulated by the law."[272] In this sense, the civil order has both a more formal, legal dimension which includes "the institutions and legally ordained practices by which it formally governs itself", and a dimension which includes the "informal, extra-legal practices and expectations by which our social lives as citizens are structured."[273] From this conception of a civil order, Duff then proposes that moral wrongdoing – specifically public wrongdoing[274] – is relevant to the criminal law as a kind of attack on the civil order of a polity. He states that "such wrongdoing can threaten civil order both instrumentally and

[270] Duff, *The Realm of Criminal Law* (n 1) 149.
[271] ibid 151.
[272] ibid 154.
[273] ibid 159.
[274] As Duff explains, "a 'public' wrong is a wrong committed in a public normative space", ibid 165. This normative space is precisely the civil order that he proposes.

intrinsically. An instrumental threat has some potential consequential impact on the functioning of the polity or its civil order, […] An intrinsic threat is damaging even if it has no such consequential impact".[275]

So, now we can return to Duff's two-part master principle for criminalisation and better understand it. By reconstructing what counts as the normative domain of the polity's civil order, we can then determine what kinds of conduct either instrumentally or intrinsically violate that order and if they do, then they can count as public wrongs. Identifying a conduct as a public wrong entails, according to Duff, that we have both a positive reason to criminalise that conduct – it is a relevant public wrong and so that gives us *a* reason to criminalise it – and a negative constraint on criminalisation based on the normative nature of the conduct – we should criminalise conduct *only if* it is some kind of public wrong. It is also, as Duff admits following Williams's terminology, a 'thin' principle in the sense that the substantive content of the principle depends on "a normative account of the civil order of the polity, and for what kinds of conduct constitute wrongful violation of that order."[276]

This particular focus that Duff gives to civil order as a central element of the framework under which criminalisation decisions are meant to be taken is an interesting point of contrast with Moore's version of legal moralism. Instead of focusing first on the properties of the conduct to potentially be criminalised, Duff wants us to focus on the normative domain within which we should be looking for properties of conduct in the first place. That being the case, a potential decision-maker's deliberation process is going to have a different starting point, or at least a different initial question: defining what is the civil order that is relevant to their decision-making. The decision-maker

[275] ibid 205.
[276] ibid 275.

will have to do this (or at least already have decided it) and only *then* move on to looking at the conduct in question, in order to verify that a) it is a conduct relevant to that domain, and b) it contains the property of the relevant kind of wrongness, i.e., it is a public wrong. Duff argues for this by pointing to how one needs to determine what "the law's business" is in order to avoid "the familiar criticism that legal moralists advocate a kind of state-controlled moralistic witch-hunt."[277] What he means by this is that more robust forms of moralism (such as Moore's) can be seen as a way by which the criminal law is used to impose moral views through the criminal law in order to impose retributive suffering on wrongdoers – a kind of pejorative sense of moralism.[278] Identifying the 'public' character of the wrongness of a conduct first, allows the decision-maker – and the criminalisation theorist – to "think about the criminal law's business, as a distinctive kind of legal institution" and, in determining said public nature, the wrong "will not need to be distinguished, within the whole realm of wrongs, from 'private' wrongs: they will come to our attention already marked as public wrongs."[279]

This difference in the deliberation process is crucial in understanding Duff's model properly. It shifts the attention of the decision-maker from the property of wrongness to the realm under which that wrongfulness appears. A large part of the potential justification or lack thereof for criminalising a conduct, under Duff's model, will depend less on whether we have correctly identified a conduct as wrongful in some sense, and more on whether we have correctly identified the *domain* (or realm, in Duff's terms) in which that wrong is found. In other words, trying to determine whether criminalising a conduct is justified will not only be determined by discovering that a conduct is morally wrong, but rather on whether that wrong is the law's business. We

[277] ibid 79.
[278] Alfred Archer, 'The Problem with Moralism' (2018) 31 Ratio 342.
[279] Duff, *The Realm of Criminal Law* (n 1) 79.

can see this difference with how Duff compares his model to Moore's model when discussing the potential criminalisation of fornication or adultery. Both can disagree with that conduct being criminalised, Duff argues, but on different grounds: "Moore can argue that [the polity] should not view consensual, adult non-marital sex as morally wrong; I can argue that we need not, and should not, discuss the morality of consensual non-marital sex in this context, since it should not be included within the public sphere".[280] Thus, that initial step of what wrongs get into the public sphere becomes crucial, because any justificatory or explanatory role that moral wrongfulness might play in a criminalisation deliberation hinges on the 'public' character of that wrong.

4.2.2.a. Applying the questions to Duff

From the previous discussion of Duff's proposed principle-based deliberation process, the conceptual question becomes particularly salient: what are the conceptual boundaries of what gets to count as the 'political morality' of the civil order? Duff's apparent answer to the conceptual question is especially important, since he proposes that it is not *all* of morality that criminal law should be interested in. Importantly, however, Duff is *not* saying that public wrongs are in a *different* normative domain than non-public moral wrongs, but rather he is arguing that the moral wrongs *become* part of the public sphere and, therefore, the wrongs become part of *political morality*. Duff himself states that his version of legal moralism is a "legal-political moralism, to make clear that the morality in question, in whose light crimes are defined as wrongs, is a political morality."[281]

[280] ibid 168.
[281] ibid 311.

But, as we have seen earlier, part of the conceptual question asks about what makes morality distinct as a kind of normative domain, and a large part of answering that question comes from defining the authoritativeness of the normativity of morality. Thus, for Duff's model, we need to reconstruct both what his answer would be for morality more generally, and for political morality more specifically.

To begin, it is clear that Duff thinks that there are correct answers to the moral status of certain conduct, particularly in view of his interest in 'critical' morality and his claim that criminalisation should be grounded on "actual rather than merely perceived wrongdoing."[282] We can also see this in his reply to dissenters regarding the criminalisation of certain conduct because they question the moral wrongfulness of the conduct in question. The example Duff gives is a man arguing for a moral right to beat his wife, to which Duff states that "the appropriate reply is to insist that such conduct is a serious wrong—and that he is wrong to deny this."[283] Note that he believes this kind of reply can be given *both* to a dissent on the 'wrong' aspect of the conduct, as well as of the 'public' aspect of the conduct.[284] This would suggest that for Duff, at least at a conceptual level, both kinds of claims – about morality more broadly and about political morality specifically – have truth values and one can be correct or incorrect in speaking about either normative context.

Borrowing from Hart's terminology,[285] Duff argues that one can criticise the criminal law of a particular polity from an internal point of view and an external point of view.[286] The former would criticise the inconsistency between the criminal law and a polity's self-defining values. The latter would

[282] ibid 54.
[283] ibid 131.
[284] See footnote 126 in ibid 131.
[285] Hart, *The Concept of Law* (n 192).
[286] Duff, *The Realm of Criminal Law* (n 1) 166.

criticise the conception of the civil order of that particular polity. Note, however, that Duff recognises that any critic will be "operating within a much messier conceptual and normative structure"[287] than what one might think at first glance. That is, civil orders are not necessarily uniform, consistent, or fixed, and the delineation of both internal and external perspectives might be more difficult to accurately determine because of that. However, Duff is also quick to accept that one can reasonably criticise a polity's civil order based on an ideal, aspirational normative model of how a polity should be –the fact there might be significant disagreement about how a polity's civil order is constituted does not exclude the possibility of criticising it from a position of an ideal that we should work towards, as "we can still recognize it as an ideal towards which we can sensibly aspire."[288]

This ideal standard for what should form the civil order's normative content is formed, for Duff, from the necessarily relational aspect of being a member of a polity. For Duff, the fact that people are part of a polity and the role they play as members of it is crucial in understanding the formation of a civil order, particularly because "their membership of the polity gives them distinctive reasons to act, and react, in certain ways towards each other—reasons that flow from *the goals and values that define their polity as a distinctive practice*; reasons that they do not have in relation to those outside the polity, and that others do not have in relation to them."[289] In this sense, it seems that Duff's view reflects a more constructivist understanding of the conceptual delimiting for political morality, since the standpoint of being a member of a polity seems to bring with it, according to Duff, certain goals and values that are distinctive of that polity. In this sense, as he points out, "an account of the *res publica* of a polity (as of any practice) will also be an

[287] ibid.
[288] ibid 182.
[289] ibid 101. Emphasis added.

account of the kinds of agent-relative (practice-relative) reasons for action that apply to its members, and that will therefore underpin and guide its laws."[290]

So, we can start to attempt constructing an answer at this point to the conceptual question. The relevant answer to this question, for Duff, has to do not only with the authoritative normativity of morality generally, but specifically with regards to the civil order of a polity and how the moral facts included in that context make a morally wrongful conduct the law's business. Thus, the authoritative normativity of this 'public' domain is going to be stance-dependent, relative to holding the position of 'member of a polity.' For those who are members of that polity, the reasons and values of the civil order are binding, and are not escapable through a desire to be rid of them. As Duff puts it, if someone disagrees with the polity's civil order and its authority over them, "the only answer (which is a perfectly adequate answer) is that they, and we, are stuck with each other as fellow members of this political association, and that the associative duties that they, and we, thereby acquire are not to be voluntarily rejected."[291] However, the conceptual understanding of what counts as 'moral' in 'public moral wrong' is different because, as we shall see in discussing the ontological question, Duff does not seem to view morality as stance-dependent.

However, as to providing an answer to the ontological question, Duff gives us very little guidance, if any at all. We can, however, infer some possible content for a reconstructed answer to the question. In particular, his treatment of *mala in se* offences can be helpful, since Duff is willing to accept that they are wrongs "because that is not how one should treat anyone, fellow

[290] ibid 100.
[291] ibid 181.

citizen or not."[292] Duff does not provide much explanation as to why this is the case independently of the polity's civil order, but if we contrast this view with his understanding of the 'public' nature of political morality as agent-relative, we might be able to get some clues. There is nothing in Duff's writing that suggests he thinks this agent-relative view also holds for moral wrongness more generally – it seems, rather, that he believes that said kind of wrongness just *is the case* regardless of any position one holds (other than, perhaps, being a human being). So, the answer to the ontological question is probably not going to go down a constructivist route for morality more generally, but it does seem to go down that route for political morality to some extent. This creates an important question regarding why this can be the case, which Duff does not really provide. As Tomlin rightly points out, for Duff "there is only one true morality, but there is reasonable pluralism when it comes to constructing the civil order."[293] In this sense, Duff's answer to the ontological question is not necessarily consistent, and though this might seem problematic at first glance, it is the kind of issue that could be, at the very least, discussed more thoroughly with tools from metaethics.

In view of this dual nature to the ontological answer, a reconstruction for the epistemic question will also need to be divided. Thus, the epistemic question is both broader and more specific in Duff's case. It does ask how we come to know the properties we are looking for to assign wrongness to a conduct, but it *also* asks how we come to know the properties of the *domain* where that wrongness is found. For the first kind of inquiry, into pre-legal moral wrongs, Duff does not provide us with explicit guidance. Instead, he seems to think that for many cases, particularly for *mala in se* crimes, the moral

[292] RA Duff, 'Responsibility, Citizenship, and Criminal Law' in RA Duff and Stuart P Green (eds), *Philosophical Foundations of Criminal Law* (Oxford University Press 2011) 139; See also Duff, *Answering for Crime - Responsibility and Liability in the Criminal Law* (n 4) 143.
[293] Patrick Tomlin, 'Duffing Up the Criminal Law?' (2019) 14 Criminal Law and Philosophy 319, 328.

wrongness of the conduct is obvious.[294] Therefore, there is little to go on to reconstruct an accurate epistemic method for broad moral wrongs, though the language of 'obvious' tends to suggest some form of intuitionism as part of the epistemic process (which, of course, does not entail a lack of scrutiny of those intuitions later through some form of reflective practice). However, the same 'obviousness' could be the result of forming a coherent belief based on a prior set of moral beliefs, thus inferring the moral wrongness as obvious rather than intuiting it.[295] Therefore, the epistemic method for moral wrongness, for Duff, remains open to interpretation.

For public wrongness that is relevant to political morality, however, the epistemic process seems slightly clearer. The first step will be to establish which values are "intrinsic" to the practice at hand, in this case, the civil order.[296] In doing so, we will need to put ourselves, as decision-makers, in the epistemic position required to gain access to both an idealised external perspective of a polity's civil order, as well as a more descriptive internal perspective of how the civil order actually manifests in practice. This will allow our deliberation to have a framework under which we can determine the relevance of a conduct's wrongfulness to the civil order, thereby obtaining its 'public' nature. So, epistemic access to that framework will require a rational reconstruction of the ideal values that are coherent with a civil order, and a further rational reconstruction of a best explanation for how the polity's civil order in particular is constructed in practice. In order to do this, a decision-maker is going to have to think about several different things, some of which will be more empirically informed than others. For example, it seems plausible that Duff would agree, in view of the fact he sees his principle as

[294] See, for example, the discussion of murder and rape as "obvious public wrongs" in Duff, *The Realm of Criminal Law* (n 1) 300–302.
[295] This is particularly plausible considering Duff's methodology of rational reconstruction, set out in ibid 11–13.
[296] Duff uses the language of "intrinsic" values at ibid 13.

thin regarding its substantive content, that a decision-maker will need to look at political theory in order to gain epistemic access to the external point of view. They will need to make theoretically-informed decisions as to which political theories provide the best values for a civil order, which might entail committing to particular political ideologies or normative frameworks. For the internal point of view, in contrast, decision-makers will need to look at other possible sources. They can, for example, seek empirical evidence as to what members of a particular polity hold to be the values of their civil order. They can also take into account things like the existing constitutional framework of a polity, or the political commitments that may have been taken on by the polity's state. Regardless of the specific sources used, however, the important thing to note is that the decision-maker will then have to use rational reconstruction, according to Duff, in order to provide themselves with the normative framework under which they can then determine the 'public' nature of the wrongdoing in question.

4.2.3. Douglas Husak – Overcriminalisation

Husak presents his theory of criminalisation as a "minimalist" criminalisation theory, which is formed by "a total of seven general principles or constraints designed to limit the authority of the state to enact penal offences".[297] He divides these principles into *external* and *internal* constraints on criminalisation — external because "they depend on a controversial normative theory"[298] of the state, and internal because they are derived "from the criminal law itself."[299] The external constraints are that the state must have a substantial interest in the statutory objective being achieved, the law must directly advance that interest, and that the statute must be no more extensive

[297] Husak, *Overcriminalization: The Limits of the Law* (n 1) 55.
[298] ibid.
[299] ibid.

than necessary to achieve its purpose.[300] The internal constraints are the nontrivial harm or evil constraint, the wrongfulness constraint, the desert constraint, and the burden of proof constraint.[301]

Note that the principles Husak proposes are framed as constraints, which does not mean that they do not play any justifying role in his theory of criminalisation, but rather that they justify by *limiting* the cases in which criminalisation will be justified. In this sense, both sets of principles are doing justificatory work for criminalisation, but their targets are different. External principles seek to "address not only the persons who are punished but also the citizens who are asked to create and maintain a system of punitive sanctions",[302] thereby focusing on the broader purposes that a state may have when criminalising and the legitimacy and justification of those purposes. Internal principles, in contrast, are addressed to those individuals who are punished, and can serve as a way to justify the application of punishment towards those individuals.[303] Importantly, Husak's theory is very much based on the idea that criminalisation involves criminal punishment. He believes that it is "the most important difference between the criminal law and other bodies of law, or between the criminal law and systems of social control that are not modes of law at all"[304] and, therefore, to him "a theory of criminalisation provides the set of conditions under which the state is permitted to resort to punishment."[305] In this sense, Husak's reasoning is similar to Moore's views on the function of the criminal law, specifically its connection to punishment, to the point that even though Husak does not share Moore's view of retributive

[300] ibid 132.
[301] ibid 55.
[302] ibid 121.
[303] ibid.
[304] ibid 77.
[305] ibid.

justice as the value being pursued by punishment,[306] he does think that "Moore understands the implications of a theory of punishment for a theory of criminalization."[307]

In this sense, Husak's internal principles are the more relevant to examine for my purposes, as they purport to justify why a conduct should be selected as one which leads to criminal punishment for the individual that carries out the criminalised conduct, and it is where Husak has explicitly placed wrongfulness as a concept to be used in theorising about criminalisation. However, the external constraints are also relevant to the discussion here, since Husak believes, in line with Duff and Marshall,[308] that the *kind* of wrong that becomes relevant to the criminal law is defined by determining the legitimate interests of the state and, therefore, which wrongs ought to be considered public wrongs.[309] In this sense, though the use of wrongfulness may be more noticeable as a decision-making principle at the level of internal constraints, the external constraints are still going to play a justificatory role in defining the relevance of the wrongs identified, and therefore on whether the decision to criminalise is ultimately justified. Husak also proposes that his external constraints are "the most important difference between my own theory of criminalisation and legal moralism".[310] He argues this because, though the legal moralist will accept his internal constraints – particularly his wrongfulness constraint – his external requirements regarding legitimate state interests and how those are achieved by criminalising are not necessarily going to be accepted by the legal moralist.[311] Husak's emphasis in using the external constraints is, again, linked to justifying punishment

[306] ibid 200–206.
[307] ibid 197.
[308] Marshall and Duff (n 260).
[309] Husak, *Overcriminalization: The Limits of the Law* (n 1) 135.
[310] ibid 198.
[311] ibid.

through criminalisation, but he argues that in adding the external constraints he is not falling into the legal moralist (*a-la*-Moore) exercise of justifying that punishment purely in retributivist terms. Rather, Husak argues, he is using the external constraints as a way to determine when the state has legitimate and appropriate reasons to pursue punishing people for doing certain acts, which *includes* but is not *limited to* the moral wrongfulness of the conduct in question. In his words, the external constraints (and his 'minimalist' criminalisation theory in view of them) allow for decision-makers to fulfil "the need to show why the *state* is the appropriate vehicle for imposing punishment."[312]

It is also at the internal constraints level where Husak has placed moral wrongfulness in his model, as a necessary (but not sufficient) condition for justifying criminal punishment on someone: "criminal liability may not be imposed unless the defendant's conduct is (in some sense) wrongful."[313] He derives this constraint, along with the other three constraints of non-trivial harm, desert and burden of proof, from "within the boundaries of criminal theory as it is presently conceptualised, even though these resources have not been explicitly utilised for this purpose."[314] That is to say, because of how the general part of the criminal law currently stands, Husak argues, one can derive these internal constraints on criminalisation as they are required to make certain doctrines, like excuse defences and justifications, intelligible.[315] Note, however, that because of the way in which Husak ends up defining 'harm' as his first constraint, under which 'non-trivial harm' and 'evil' can fall into the category of harm,[316] it is arguable that harm ends up collapsing into the wrongfulness constraint, rendering the harm constraint unnecessary.

[312] ibid 200.
[313] ibid 66.
[314] ibid 57.
[315] See ibid 66–77.
[316] ibid 70 and n70.

Indeed, as Moore points out, "conjoining the second [wrongfulness] constraint to [the first constraint] yields: behavior must be wrongful (second constraint) *and* (from the first constraint) either harmful *or* evil, where 'evil' is read as harmless but wrongful. Translation: behavior must be wrongful or harmful to be prohibited."[317] In this sense, Husak's position ends up being closer to Moore's legal moralism than he might have anticipated.

From the importance and justification of criminal punishment, and the derived principles from the general part of the criminal law, Husak arrives at this summarised version of his criminalisation principle (based on internal constraints): "Penal statutes must proscribe a nontrivial harm or evil; hardship and stigma may be imposed only for conduct that is in some sense wrongful; violations of criminal laws must result in punishments that are deserved; and the burden of proof should be placed on those who advocate the imposition of criminal sanctions."[318]

A reconstruction of Husak's deliberation process, then, would look something like the following. A criminalisation decision-maker will have to begin their deliberation by identifying whether a candidate conduct includes non-trivial harm or evil and make sure that their proscription is purposefully targeting that harm. After doing this, wrongfulness *then* enters the deliberation process as a constraint, leading the decision-maker to ask themselves whether the harm they have identified is wrongful or not, and if it is so in the relevant sense. Wrongfulness, as Husak presents it in his model, has two functions: as a way to identify the harm being targeted as relevant to the deliberation; and as a necessary justification for punishment, which links it to Husak's third desert constraint. The decision-maker will need to,

[317] Moore, 'A Tale of Two Theories' (n 179) 39. Emphasis in the original.
[318] Husak, *Overcriminalization: The Limits of the Law* (n 1) 103.

therefore, explain how the wrongness they have identified can serve as part of the justification for applying punishment to offenders, which Husak embodies in the idea that punishment must be *deserved*, and the way in which that desert is identified is through the presence of the appropriate wrongfulness in the conduct being punished. Finally, the decision-maker will have to bear the burden of this justification exercise, according to Husak's fourth internal constraint, because "punishments implicate and potentially violate important rights—the right not to be deliberately subjected to hard treatment and censure by the state".[319] So, the deliberation that Husak proposes will require the decision-maker to be able to provide said justification, precisely by alluding to the wrongness of the conduct, and explaining how that wrongness leads to deserving punishment.

As for the specific ways in which Husak presents the nature of wrongdoing, there is not much explicit reference to it. In fact, Husak admits that "the real challenge for legal philosophers is not to demonstrate the bare existence of the wrongfulness constraint, but to infuse it with substantive content" and states that he wishes to stay as ecumenical as possible, appealing to intuitions that "involve familiar, everyday situations".[320] There are, however, some ideas that he establishes around his views on the way wrongdoing works. He does states that wrongfulness must be so in *some* sense,[321] which suggests that there might be space for different senses in which a conduct can be considered wrongful, though it is not clear whether he means 'senses' as in belonging to different normative contexts or something different. But, more needs to be said because, as Tadros rightly puts it, "without an independent theory of what is wrong, it is difficult to know what is constrained."[322] Though

[319] ibid 100.
[320] ibid 76.
[321] ibid.
[322] Victor Tadros, 'The Architecture of Criminalization' (2009) 28 Criminal Justice Ethics 74, 77.

Tadros's point in this quote is regarding a lack of first-order moral theory on what is morally wrong, I submit that the same criticism can hold regarding second-order metaethical theory of what gets to count as the target wrongfulness in the first place.

Husak also makes an interesting point regarding the definition of the boundaries between the 'public' and the 'private' spheres, where he states that he is sceptical of being able to adequately define said boundary on the basis of the target conduct. In his words, "I believe that it is misguided to suppose we might generate a list of specific behaviours that are inherently private, immune for all time from state punishment. Change in social circumstances can shift the boundary between the public and private spheres."[323] He does, however, offer a partial solution for delimiting public and private wrongs specifically, at a given time and place, "by trying to determine whether the individual victim or the community should control the decision to initiate and pursue a complaint."[324] In doing so, Husak is trying to highlight the fact that, as he puts it, a theory of criminal responsibility "must be *relational* in specifying the person or body *to whom* we are responsible."[325] What is interesting, however, is that Husak leaves the door open for this relation to change over time, and that the conceptual delimitation of what gets to count as part of the public sphere is context-sensitive. What is also interesting about Husak's use of Duff and Marshall's model of public wrongs is that he does not view it as an expressive function that the state might have to express wrongdoing to the public, as Duff and Marshall do,[326] but instead he uses it as a boundary-setting for his external constraints on criminalisation, as the way in which one can identify the wrongs that are relevant to the state

[323] Husak, *Overcriminalization: The Limits of the Law* (n 1) 135.
[324] ibid 136.
[325] ibid. Emphasis in the original.
[326] Marshall and Duff (n 260).

and its decision to punish people for that wrongdoing.[327]

4.2.3.a. Applying the metaethical questions to Husak

From this statement of Husak's minimalist criminalisation theory, let us now try to reconstruct the potential answers to the metaethical questions that could fit his theory, starting with the conceptual question. First, as stated earlier, it is worth noting that Husak is portraying his wrongfulness constraint as requiring wrongfulness *in some sense*, which does not mean that he is necessarily thinking of moral wrongfulness specifically. However, there are some additional facts that might help us read his use of "wrongfulness" as referring specifically to moral wrongfulness. The first is Husak's use of desert as part of the constraints on criminalisation. The language of desert is, typically, a marker of morality as the normative domain that is being discussed. This is particularly true in the sense that desert is usually associated to blaming – a morally wrongful action will entitle the wronged party to blame the wrongdoer for their wrongdoing, *because* they deserve that blame.[328] And, as discussed above, the way in which one is supposed to justify that desert is by correctly identifying the wrongness in the conduct being judged. So, if the wrongness that Husak uses as a constraint on criminalisation is linked thusly to desert, then that wrongness is almost certainly going to be of a moral kind.[329] Husak himself points to this moral character elsewhere, stating that "impositions of punishment, and even threats of punishment, require a *moral* justification."[330] Presumably, since Husak uses a wrongfulnsss

[327] Husak, *Overcriminalization: The Limits of the Law* (n 1) 136–137.
[328] Husak himself argues for this in Douglas Husak, 'Wrongs, Crimes, and Criminalization' (2019) 13 Criminal Law and Philosophy 393, 401–402.
[329] I say "almost certainly" because, strictly speaking, there might be an argument to be made that one might "deserve" certain appropriate reactions in other normative contexts, but this argument is outside the scope of the discussion presented here.
[330] Husak, 'Wrongs, Crimes, and Criminalization' (n 328) 401. Emphasis added.

constraint as part of the justification of applying punishment through criminalisation, this at least suggests that said wrongfulness is going to be moral in nature. He also openly states elsewhere that he has "sympathy with the general project of integrating the criminal law with moral philosophy",[331] which would suggest that his uses of the normative language of 'wrongness' and 'desert' here are, almost certainly, in a moral sense. That being the case, if part of that moral justification required using wrongfulness, it would seem strange to think that said wrongfulness is not also of a moral kind.

The second clue is that Husak accepts, for his purposes, the definition of a *malum prohibitum* offence as "when the conduct proscribed is not wrongful prior to or independent of law".[332] He accepts this view precisely to then argue that most attempts at squaring *mala prohibita* offences with the wrongfulness constraint fail, because "no one has shown why the good reasons to discourage the conduct proscribed by many *mala prohibita* offences justify state impositions of hard treatment and stigma."[333] I argue that it is possible to read Husak here as saying that the wrongfulness that is prior to or independent of law and that allows justifying the imposition of hard treatment, which is present in *mala in se* offences, is of a moral kind. If the failure of squaring *mala prohibita* offences with the wrongness constraint lies in being able to justify punishment for those conducts – and as Husak argues, punishment can only be justified by alluding to *deserving* the punishment *based on* the wrongness of the conduct – then the wrongness that *does* justify punishment must be, for Husak, of a moral kind. This does not mean that one could not imagine other kinds of wrongness that can exist prior to the law more generally, but what I am suggesting is that the way 'wrongness' is being used by Husak in this context must mean 'moral wrongness'. This reading also

[331] Douglas Husak, *The Philosophy of Criminal Law* (Oxford University Press 2010) 22.
[332] Husak, *Overcriminalization: The Limits of the Law* (n 1) 105.
[333] ibid 119.

makes sense when we see that he is arguing in favour of a "minimalist criminal law",[334] under which *mala prohibita* offences are mostly relegated to being excluded from the criminal law, as they cannot fulfil the wrongfulness constraint that Husak proposes. Thus, at least for Husak, part of that fulfilment of the constraint has to do with the fact that the potential wrongness one could find for certain *mala prohibita* offences is *not* of a moral kind. Additionally, Husak argues that the offences that *do* fulfil his proposed constraint do so in a way that many existing *mala prohibita* offences do not, and he states that "I fail to understand why persons behave *wrongfully* when their conduct is *malum prohibitum* but not *malum in se*."[335] Thus, again, it seems that his use of 'wrongful' in this context is to point to specifically moral wrongness.

If we accept this reading, the problem then becomes trying to determine which conceptual understanding of 'moral' makes more sense for Husak's constraint. We need to determine both whether Husak sees morality as an authoritatively normative domain, and what he proposes (if at all) is the particular nature of authoritative normativity regarding the linking of particular properties to values. Regarding both parts of the conceptual question, Husak remains mostly silent in Overcriminalization. Elsewhere, we can get some clues as to what the answer might be, based on Husak's views on how desert might or might not be affected by the presence of a legal norm.[336] When discussing this, Husak speaks of an "all-things-considered justification" for punishment, and how it is not sufficient to use desert as the basis for establishing that kind of justification.[337] That is, Husak argues that though desert is a reason-giving property for how agents are supposed to be

[334] ibid.
[335] ibid 112. Emphasis added.
[336] Philip Mullock, 'The Logic of Entrapment' (1985) 46 University of Pittsburgh Law Review 739, 387–391.
[337] ibid 389.

treated, and culpable wrongdoing is a good candidate for this kind of reason, the state might still lack an all-things-considered justification for punishing that culpable wrongdoing.[338]

That being the case, a similar argument about the moral nature of this justification can be made to the one I presented earlier about the moral character of the wrongness involved in desert. If the justification for punishment is linked to desert, and desert is of a moral kind, then than justification will also be of a moral kind. He also specifically says that he is interested in discussing *moral* desert, in contrast to *legal* desert.[339] This lets us know that the moral character of desert is clear, and that the fact that said desert does not provide an all-things-considered justification to punish culpable wrongdoers, suggests that Husak does not necessarily consider moral justification to be equivalent to an all-things-considered justification. Though this does not give us any definitive clues as to whether morality is characteristically an all-things-considered ought or not, it does give us an idea that, at the very least, Husak is open to the idea that a moral fact is not going to necessarily going to provide an all-things-considered reason for or against an act. So, in absence of any further material from Husak from which to construct, the constructed answer to the first aspect of the conceptual question is, at best, tentative.

As to the characteristics of authoritative normativity, specifically regarding Husak's use of wrongfulness, things get even more difficult to construct. He explicitly tries to avoid giving substantive content to his conceptual use of wrongfulness. This has already been identified as problematic by Farmer for different reasons, regarding the assumptions being made about the

[338] ibid 389.
[339] ibid 391–392.

significance of 'obvious' moral wrongs to a system of criminal law.[340] The question I am asking here, however, is even more basic: what clues do we get from Husak about what makes that moral wrongness 'obvious'? The only clue we have is based on what Husak does *not* say. That is, nothing in what Husak says suggests that he is thinking of moral claims as only formally normative. Though this is obviously not a definitive indicator of Husak's view, it can at least be said that if moral claims are not seen only formally normative, but yet still normative, then it is highly likely that they will be seen as authoritatively normative. And, again, the only other clue we can find is that authoritative normativity, for Husak, will probably not mean necessarily that having an authoritative reason will be equivalent to an all-things-considered reason.

As for the epistemic question, again Husak does not give much guidance in the book. However, elsewhere he does make quite an open statement about what he thinks is the appropriate epistemic process under which coercive government policies can be justified. I will present the quotation in full, so as to get the complete picture of Husak's view:

> "Coercive government policies can be justified in general only by good reasons, and good reasons are constituted only by true propositions. A person is justified in believing a proposition true when the information available to her provides her with good epistemic reasons to believe it and these reasons outweigh the good epistemic reasons to doubt it that are also provided by the information available to her. Being justified, however, in believing a proposition that supports a policy does not suffice to justify one in supporting it. One must also be justified in believing that the reasons in support of this policy outweigh the reasons against it. In my opinion, one is justified in believing a judgment of this kind when it withstands one's critical scrutiny: when this judgment about the relative weight of reasons continues to seem

[340] Lindsay Farmer, 'Criminal Wrongs in Historical Perspective' in Duff and others (eds), *The Boundaries of the Criminal Law* (Oxford University Press 2010) 219–221.

> correct over time and to cohere with the other judgments one makes about the relative weight of reasons. Suppose, then, that a person is justified in believing a proposition and is also justified in believing that this proposition constitutes a reason in favor of a coercive policy that has greater weight than any reason against it. Does this suffice to justify this person in supporting this policy? We typically think that the more burdensome a coercive policy is, the stronger the epistemic case must be in favor of believing the propositions that count in favor of this policy. If a policy is very burdensome, then the epistemic grounds for believing at least one of the propositions that provide a sufficient reason for this policy must be very strong. If a policy is not so burdensome, then these epistemic grounds need not be so strong. I accept this common-sense view."[341]

Let us analyse this statement in order to reconstruct what an answer to the epistemic question might look like. There seem to be two different senses for the concept of 'justification' in this excerpt, which is particularly noticeable when Husak says that one can be 'justified' in believing a proposition in favour of a coercive policy and still not be 'justified' in supporting it. This second iteration of 'justified' seems to no longer be talking of *epistemic* justification, but rather of *moral* justification. In particular, this seems the case because Husak elaborates on how this second justification is obtained by saying that one must be 'justified' in the judgement one makes about how the reasons for and against the proposition are weighed. In other words, there is one sense of 'justification' which has to do with believing a proposition in support of a coercive policy to be true – an epistemic sense of justification – and another sense of 'justification' that relates to the *act* of supporting the proposition – a moral sense of justification. This latter sense is achieved, according to Husak, by reflection and critical scrutiny of the reasons (in the reading being offered here, *moral* reasons) one has to support or not the

[341] Douglas Husak and Peter de Marneffe, *The Legalization of Drugs* (Cambridge University Press 2005) 190–191.

coercive policy, which is distinct from the reflective process by which one evaluates whether a proposition in favour or against a coercive policy is true or not.

That being the case, it seems that Husak is seeing the moral epistemic methodology to find this second, moral sense of justification to be something along the lines of a critical reflective equilibrium: the justification for support of a coercive policy is found by critically weighing up reasons, and since this justification is of a moral kind, then it is fair to say that the same methodology could be extended to determining moral wrongness. So, extrapolating from what Husak says about moral justification, one can say that a criminalisation decision-maker is going to be justified in believing that a particular conduct is morally wrong by critically scrutinising the relative weight of reasons for and against the conduct in question. The difficulty, however, is that Husak does not give us much information as to *what* those reasons are, *where* they are found or *how* is that weighting of reasons supposed to happen in order to determine the wrongness of the conduct.

4.2.4. Andrew Simester and Andreas Von Hirsch – Crimes, Harms and Wrongs

As has been partly discussed earlier in the thesis, in "Crimes, Harms and Wrongs", Simester and Von Hirsch develop a particular view of what is happening when a conduct is criminalised and from this understanding, go on to propose a normative framework within which criminalisation can be justified. That is, from the fact that, as they put it, "the state declares that Φing is morally wrongful"[342] they then try to determine what justifies[343] the state

[342] Simester and Von Hirsch (n 1) 6.
[343] They state that their interest is in the moral justification of criminalisation, rather than whether it is lawful or constitutionally justified. See ibid 31–32.

making this declaration.

They start off by arguing that if the state is declaring a conduct to be morally wrongful – and they believe it is – then it is important that "the state should get it right".[344] That is, there is a truth-constraint on the declaration – in my terms, the assertion – that the state makes about the conduct being criminalised, and therefore the labelling of a conduct as wrongful should be done only when the conduct is *actually* wrongful. They frame this as a need derived from desert, regarding the consequences in treatment that a criminal conviction carries, where "if a person does not deserve such treatment, then she has a right that it not occur; and neither her conviction nor her punishment can be justified by such consequential considerations as deterrence."[345] In this sense, establishing the wrongfulness of an action, as they put it, "is indispensable to justifying its prohibition."[346] And this justification, as they put it, can be seen from two perspectives: an *ex ante* perspective, where the wrongness plays a role in respecting the responsible agency of the addressees of criminalisation, who can choose whether to comply with the prohibition and evaluate the conditional threatened consequence;[347] and an *ex post* perspective, where the wrongness justifies the imposition of punishment as a result of deserving the consequence in view of wrongdoing.[348] So it is clear to see that for the authors, the role that moral wrongfulness initially plays has to do with the consequences of criminalisation, namely that punishment and condemnation is at the end of the road, and that said punishment and condemnation should only be imposed on those who deserve it.

[344] ibid 19.
[345] ibid.
[346] ibid 21.
[347] ibid 8.
[348] ibid.

Once they have established this general justificatory role, they argue for three specific theses regarding it: a thesis of insufficiency (that Φ is wrongful is insufficient to justify criminalisation); a necessity thesis (that Φ is wrongful is necessary to justify criminalisation); and a non-qualifying thesis (that Φ is wrongful is insufficient for even a *pro tanto* ground for criminalisation).[349] They claim that the first thesis of insufficiency is almost trivially true, in the sense that it is generally accepted that wrongness can be outweighed and defeated all-things-considered, and defend the necessity thesis as the ultimate justificatory role that moral wrongfulness should play: any prohibition of Φing can be justified only when Φing is a morally wrongful action.[350] They defend this claim by appealing to the "distinctive nature of the criminal law, which punishes, and censures, the offender for his conduct."[351] This defence hinges on the idea that this distinctive nature carries with it the act of blaming, which is legitimate only in relation to moral wrongs. They define moral wrongs as the fact that an agent "does something that, all things considered, she ought not to do."[352] This gives us a good indication as to what kind of normative domain the authors believe morality to be: it is the realm of what ought or ought not to be done, all things considered.

Importantly, however, Simester and Von Hirsch are proposing the use of moral wrongfulness in the deliberative stage only as a negative constraint on criminalisation, rather than as a positive reason in favour of it. They claim that the third non-qualifying thesis (that moral wrongfulness does not generate *pro tanto* reasons for criminalisation) is true, and they base this claim on a further moral argument: "while an act of wrongdoing may by itself warrant censure *ex post*, something more is required to justify the *ex ante* step of

[349] ibid 22.
[350] ibid 23.
[351] ibid.
[352] ibid.

criminal prohibition. That your Φing would be wrongful doesn't entitle me to *coerce* you not to do it."[353] So, in view of their earlier claim on the dual perspectives regarding how wrongfulness can justify a criminal prohibition based on punishment at the end of the road, the authors recognise that the *ex ante* perspective cannot be fulfilled merely by stating that moral wrongfulness is present; hence why moral wrongfulness, for Simester and Von Hirsch, can only play a constraining role, rather than a positive reason role. The 'something more' that is needed, they argue, is based on the "activity at issue [being] sufficiently problematic to warrant intervention."[354] This problematic nature of the activity is established by determining that the conduct "leads to significant levels of harm and/or offence suffered by others, although there may also be limited occasions for paternalistic intervention."[355] In this sense, wrongfulness is not supposed to play a positive role in Simester and Von Hirsch's deliberation, but rather as (the most important) mediating consideration when deliberating about criminalising conduct.

This non-qualifying thesis is an important point of contrast between Simester and Von Hirsch's theory and the other theories I have presented thus far: not only because wrongfulness is seen only as a constraint, but because this changes the way the deliberation process ends up looking. This is for two reasons: first, the starting point of the deliberation shifts, and second the factor that determines relevance for the deliberation also shifts. For theorists like Moore,[356] the starting point in the deliberation process is identifying the moral wrongness of the conduct in question. For Simester and Von Hirsch, in

[353] ibid 29. Emphasis on the original.
[354] ibid 189.
[355] ibid 189. See also parts II-IV for their explanation of why harm and offence principles can serve as positive reasons in favour of criminalisation, as well as when paternalistic intervention is exceptionally legitimate.
[356] See Moore, *Placing Blame - A General Theory of the Criminal Law* (n 1) 70 and 645–647.

contrast, the starting point is one of their factors that determines the 'sufficient problematic' nature of the conduct, such as harm or offence to others. Further, once the initial features of the conduct are identified, the decision-maker needs additional elements to decide whether those features are relevant to the kind of decision they are making, in this case criminalising a conduct. For Duff, for example, this factor or relevance is the 'public' nature of the wrongfulness.[357] For Simester and Von Hirsch, rather, the moral wrongfulness of the conduct is what ends up playing this role: the decision-maker will know that the harmful or offensive conduct is of the relevant kind to criminalisation if, and only if, it can be judged as morally wrongful; hence the non-qualifying thesis. And, for the authors, the need for identifying this moral wrongfulness is linked to the fact that the criminal law has a condemnatory function, i.e. that blaming is involved. As they put it, "any formal judgment of blame, including a criminal conviction, should be predicated upon norm-violating conduct—like all blaming verdicts, it expresses moral reproof of a person *for* that norm-violating conduct."[358] Thus, even though moral wrongness is not acting as a positive reason in favour of criminalising conduct, it does play a central role in the deliberation process for criminalisation to determine when it is appropriate for the criminal law to fulfil its role as a formal judgment of blame.

They also base the claim for the non-qualifying thesis on their understanding of the role of the state, whereby it is only an instrumental actor to advance the welfare of its subjects and, based on this, moral wrongfulness should not be the starting point for criminalisation. Instead, they argue, "The starting point […] is the lives of people. The state has an interest in regulating

[357] As he puts it, "to count a wrong as 'public' is precisely to assert that it is, in principle, the law's business—rather than a substantive ground or criterion for that conclusion" in Duff, *The Realm of Criminal Law* (n 1) 78–79.
[358] Simester and Von Hirsch (n 1) 71. Emphasis in original.

conduct that affects those lives. And where such conduct is wrongful, it may be appropriate to regulate it by the criminal law."[359]

Interestingly, their claim for the necessity thesis as the justificatory role for moral wrongfulness is broader than for only the context of criminalisation. They claim that it applies to "all of us", including the state as a kind of moral agent who is held to the same general obligation not to act wrongfully. As they put it:

> "the state is, like the rest of us, subject to requirements of morality. This matters because unless, all things considered, it is wrongful for D to Φ, it generally follows that third parties—including the state—should not object to D's Φing; let alone prevent it. Notwithstanding that D's reason (or duty) to Φ may be personal to D, the existence of those reasons is itself a general matter. That D has reason to Φ, therefore, is something that commands our allegiance too. This is not so much because D herself is entitled to respect—since respect for another human being does not imply that we must always respect the reasons for which that person acts. Rather, it is because the reasons themselves are entitled to respect—that is, because they are (good) reasons."[360]

So, we can now see the outlines of what the deliberation process looks like. The criminalisation decision-maker will have to carry out a deliberation under which they are able to identify if the conduct in question is, all things considered, wrongful (and, thereby, morally wrongful). But, importantly, this moral wrongness does not provide even a pro tanto reason in favour of criminalising that conduct. Instead, the fact that the conduct is, all-things-considered, a morally wrong thing to do can only act as a mediating constraint for the criminalisation decision. In a sense, using moral wrongfulness in this way acts as a marker of relevance for the decision-maker

[359] ibid 30.
[360] ibid 24.

when deliberating – if you see a conduct that is sufficiently problematic (i.e. is harmful or offensive in the relevant way), you need to then check if that conduct is, all things considered, wrong. If it is, then the decision-maker can be sure that it is the right kind of conduct to be considering for criminalisation.

4.2.4.a. Applying the questions to Simester and Von Hirsch

As has already been stated, the answer to the conceptual question the authors would give is relatively straight-forward. For them, moral wrongs are actions that, all things considered, one ought not to do. This definition is not very helpful in order to identify the specific properties of the conduct that constitute moral wrongness, but we can still see one way in which said wrongness can be identified: it will be the result of a deliberation of reasons and, once all relevant reasons have been considered for and against the action, the reasons against doing the action will win out, making the action, all things considered, the wrong thing to do. The authors equate this to 'moral' wrongness, and so their potential answer to the conceptual question will be something along the following lines: moral wrongness is equivalent to all-things-considered wrongness.

There are some further clues as to what their answer might be to the conceptual question. In their treatment of the offence principle, they argue that the appropriate way to use offence as a part of the criminalisation deliberation is by joining it with moral wrongness. As they state it, "it is not affronting the sensibilities of other persons (even many of them) that should justify possible state intervention, but affront *plus* valid normative reasons for objecting to the conduct."[361] This suggests that the fact that a person (or group of people) experience an affront to sensibilities, is not equivalent to

[361] ibid 96. Emphasis in original.

there being moral wrongness, thereby excluding any desire-dependent accounts of moral wrongs. Notice as well that the language of 'valid normative reasons' as being *added* to the affront or offence suggests that the authors believe that the mental state of offence, which is caused by the conduct, is *not* what provides the wrongfulness *to* the conduct but rather it is this 'valid normative reason' that does so. In this sense, we can also see why, as stated above, moral wrongness ends up acting as a kind of marker of relevance for the decision to criminalise, since it allows for the decision-maker to track conduct that can appropriately be censured by the state through the criminal law – this is specifically done by identifying moral wrongness, since it is this kind of wrongness that merits, according to the authors, censure and blaming. In the context of offence, for example, the authors say that "with conduct that supposedly is offensive, one must thus ask why the actor has done anything to deserve censure. If the gravamen of offence is merely that the conduct displeases many people, then it is not clear that wrongdoing has occurred at all. Better reasons need to be provided why and in what respects there is something reprehensible about the behaviour that properly may attract a censuring response."[362] And this additional 'better' reason, it is fair to assume, is what the 'all-things-considered' nature of moral wrongness would bring to the deliberation process.

We have at least some basic clarities as to what the deliberation process for criminalisation, according to Simester and Von Hirsch, looks like. But do the authors explain the prior deliberation process that a decision-maker needs to do to figure out that all-things-considered wrongfulness? Not particularly. This is where metaethics will become a relevant consideration, particularly regarding what different views one can believe 'all-things-considered' entails. Since the authors characterise moral wrongness as what,

[362] ibid 97.

all-things-considered, an actor ought not to do, we will need to determine what exactly is meant by 'all-things-considered' in this context, and what the all-things-considered ought looks like.

However, doing so is not a simple task. Holding this kind of position of morality as the 'all-things-considered ought' brings with it some difficult metaethical commitments to accept without question. Baker has explored these convincingly,[363] particularly regarding the ambiguity of what 'all-things-considered' might mean in this context – either all the things and facts in the world that can be considered in the decision, or all the things that *ought to be* considered.[364] Since the former meaning leads to a preposterous conclusion – that the literal quantity of reasons is what gives the all-things-considered reason its normative authority – we are left with the latter meaning of what ought to be considered. As one might imagine, this leads to a dangerous level of circularity: we need to make the normative claim that things ought to be considered a certain way, in order to *then* say that one ought to do what that deliberation has resulted in, as it is all-things-that-ought-to-be-considered what one ought to do. And trying to determine what "ought-to-be-considered" in this context is not only difficult, but is also a further normative claim that, as we might imagine by this point, is also going to be affected by one's metaethical commitments. There is also the added difficulty, based on a point made by Baker,[365] that if an "all-things-considered" ought, understood as an ought *simpliciter*, is equivalent to a moral ought, then that leads to statements like "the killer ought to have covered his tracks better" to be non-sensical, as this would probably not be an acceptable statement of a moral ought. And yet, that statement is intelligible and can be

[363] Derek Baker, 'Skepticism About Ought Simpliciter' (2018) 13 Oxford Studies in Metaethics.
[364] ibid 237.
[365] ibid 250.

compatible with saying that under one "all-things-considered" perspective – say, self-interest – the killer *really* ought to have covered his tracks better, and at the same time believe that, morally speaking from an "all-things-considered" perspective, the killer ought to confess to their crimes.

So, we are left with a potential answer to the conceptual question which, at the very least, can be reasonably questioned from a metaethical perspective. As for the ontological and epistemic questions, the book does not provide much more clues as to possible answers to be constructed. However, there is one issue that the authors discuss which can at least suggest a potential metaethical commitment regarding moral wrongness as something more basic or obvious, in contrast to other evaluative standards. And, importantly, they seem to think that the meaning of 'morally wrong' is both distinct and easier to determine than other terms like 'reasonable' and 'dangerous'.[366] They say that, as part of the considerations around fair warning, evaluative terms are "imprecise, and for that reason, at least in serious crimes, they should normally be used only sparingly and for clear cases of wrongdoing."[367] One reading of this is a suggestion that there can be cases where the moral wrongness of the conduct (at least in serious crimes) is clear cut, whereas other evaluative aspects of it are not. Thus, this would imply that the ontology of moral wrongness and the epistemic method by which it can be found will be, to at least some degree, accessible to most people, including potential criminalisation decision-makers. This is inconclusive as to *what* makes moral facts like wrongness accessible, but we can combine this idea of accessibility with their view of morality as "all-things-considered" and see that the epistemic process is going to involve an accessible *reasoning* process by which we can correctly identify what makes

[366] Simester and Von Hirsch (n 1) 200.
[367] ibid 201.

it the case that, all-things-considered, an agent ought not to do the criminalised conduct.

Finally, we need to also remind ourselves of Simester and Von Hirsch's point on the truth-constraint for condemnation based on moral wrongness.[368] Since the authors believe that the state can "get it right" regarding the truth-tracking of a statement of moral wrongfulness, this suggests that the authors are accepting a commitment to truth-aptness of moral facts – if the state can "get it wrong", that entails there being a true statement of moral fact that the state is not capturing. Thus, I propose the following as a tentative construction of the answers that would be given by the authors to the metaethical questions posed: moral wrongness is a truth-apt fact which is determined by identifying reasons regarding the conduct in question and, through a rational deliberation, a decision-maker can determine that it is the case that the conduct, all-things-considered, ought not be done. Again, this is where having a clearer discussion by the authors of the metaethics involved in moral thought would be useful.

4.2.5. Victor Tadros –Wrongs and Crimes

In his contribution to the Criminalisation series of books, Victor Tadros's monograph Wrongs and Crimes is perhaps one of the most nuanced and complex studies of the theorising of criminalisation. As the title suggests, a large part of the text is focused on analysing the relationship between moral wrongs and crimes, and in trying to understand why they are, supposedly, linked in more than just a contingent way.[369] He starts from the intuitive idea that moral wrongfulness is linked to criminal justice in view of the existence

[368] ibid 19.
[369] See Tadros, *Wrongs and Crimes* (n 1) 11.

of punishment – "the central problem of punishment is how to justify a systematic state practice of intentional harm infliction."[370] To do this, Tadros argues, one must take a closer look at the importance of wrongdoing: at why it is this, and not something else, that makes a person liable to punishment.[371]

In order to understand this, he proposes that we must first understand what is meant by moral wrongdoing and what role it plays in our practical deliberations. By moral wrongs he means a distinctive kind of wrong, that is different from, for example, wrongs against taste.[372] So, for Tadros, what makes moral wrongness distinctive becomes an important issue to determine before even beginning to use it (and evaluate its use) within criminalisation theory. At its core, this distinctiveness for Tadros has to do with the fact that moral wrongness creates a "decisive reason" against performing some action.[373] This decisive reason makes us "unfree" to perform that action and, therefore, "wrongdoing constrains—these constraints naturally follow from a recognition that it would be inappropriate to value the free performance of an action as an expression of one's agency."[374] However, Tadros then goes on to add more distinctive features, and focuses on the *normativity* of moral wrongness. For example, he states that "the normative force of moral wrongness is not simply the moral force of having a decisive reason; it is normatively compelling in a way that mere decisive reasons are not."[375]

Tadros links this normative force to a particular kind of affective reaction to the presence of moral wrongness. As he puts it, "the normative force of

[370] ibid 12.
[371] ibid 13.
[372] ibid 16.
[373] Tadros makes an important distinction between 'wrongness' and 'mere irrationality' which escapes the scope of this section. See ibid 15–25.
[374] ibid 42.
[375] ibid 28.

morality explains our reaction to acting morally wrongly, when we recognize that we have. We reflect on the decision to act wrongly. We now feel the normative force of the facts that make the act wrong: something that we either failed to recognize or that failed to move us."[376] This is precisely what he means when he says that "when an act is wrong, we do not simply have a very weighty reason not to do it; we are constrained",[377] and that "morality constrains the pursuit of the things that we value most, including things that we value most for good reason."[378] This is because, as he puts it, "moral wrongness makes an act compelling in a distinctive way", and importantly "the way that wrongness is compelling is normative rather than psychological: when an act is wrong a person is compelled not to perform it in virtue of the reasons against performing the action."[379]

Notice that this is not the same as saying that the normativity that distinguishes moral wrongness is *response*-dependent, as in saying that something is morally wrong because it is appropriate to blame others for their wrongdoing, which is a position that Tadros rejects.[380] Instead, the point is regarding the affective reaction that the *agent* feels towards the recognition of moral wrongdoing, and particularly in saying that the normative force of moral wrongness is what causes that reaction, rather than the reaction being what explains moral wrongness.[381]

This distinctness, however, still shares commonalities with other kinds of wrongdoing, which Tadros explains by stating that "if it is wrong to *v*, *v*ing is

[376] ibid.
[377] ibid 39.
[378] ibid 32.
[379] ibid.
[380] ibid 17–25.
[381] See also Victor Tadros, *The Ends of Harm - The Moral Foundations of Criminal Law* (Oxford University Press 2011) 47–48.

ruled out. This character of wrongness implies that we can express wrongness claims as requirements of a kind" and then goes on to say that "an account of moral wrongness aims to characterise the distinctive kind of requirement involved in moral wrongness."[382] I will discuss what Tadros says about this distinctness in the next sub-section.

From there, Tadros eventually goes on to present a deliberation process for criminalisation regarding serious wrongs.[383] He does not endorse this process as the correct one for all criminalisation decisions, but rather states it as the argumentative structure that would be needed in order to make moral wrongfulness play a role in the criminalising of serious wrongs. He states that we must first identify a conduct as wrong, and explain why the conduct is wrong, for which we will need a definition of the nature of wrongdoing and to identify the source of wrongdoing in particular.[384] We then need to explain why the state is warranted to create an enforceable prohibition on the conduct, and show that the punishing of those who carry out the conduct is permissible.[385] To do the latter, one needs to show that the violation of a prohibition creates liability to punishment, and show that the liability is grounded in duties of offenders, in the sense that the wrongdoer has a duty to respond to the wrongdoing, and that the duty is related to the ends of the criminal law.[386] Then, Tadros adds a requirement to show that criminalising the conduct in question is proportional, and he divides proportionality into two senses: a narrow sense, under which punishment is proportionate to the stringency of the duties that offenders incur through their wrongdoing; and a wide sense, under which the bad effects of criminalising the conduct are

[382] Tadros, *Wrongs and Crimes* (n 1) 16.
[383] See Ch. 9 in Tadros, *Wrongs and Crimes* (n 1)., with a summary of the process in 172
[384] ibid 159–162.
[385] ibid 162–166.
[386] ibid 166–170.

outweighed by the good that criminalisation achieves.[387]

The first and second steps are key to the whole process if using wrongness is to work, since as Tadros puts it, for this kind of deliberation process "the central reasons to prevent the [criminalised] conduct have to do with the facts that make it wrong: facts that we have independent reasons not to want to occur."[388] It is also important to note that, although Tadros himself rejects the use of moral wrongfulness as the main positive reason for criminalisation based on reasons of desert like Moore,[389] he recognises a core of correctness in the argument for wrongfulness as a negative constraint (i.e. a necessary condition), in the sense that, as he puts it, "the fact that a conduct is wrong […] plays a negative role in judgments about criminalisation: certain reasons that sometimes count against criminalisation do not count against it in the case of serious wrongdoing."[390] Interestingly, however, Tadros also believes that moral wrongdoing has at least some role as a positive reason for criminalisation, which has to do with the functions of the criminal law:

"The criminal law has a blaming function. There are positive reasons to communicate about wrongdoing. And this provides reasons in favour of criminalising it. A society that fails to condemn the serious wrongs that are at the heart of the criminal law fails in its moral duty to respond to these wrongs by engaging with offenders, victims, and the public at large."[391]

Tadros supports this idea by arguing that one of the additional functions of the criminal law, aside from blaming, is to express the moral duties citizens

[387] ibid 170–172.
[388] ibid 170.
[389] ibid 115–120.
[390] ibid 170.
[391] ibid.

are under by prohibiting morally wrongful acts.[392] By expressing these duties, the criminal law is fulfilling a preventive function, which is not only through some kind of deterrence in view of potential criminal punishment. According to Tadros, the criminal law "not only offers threats, it expresses to people the reasons that they have not to perform certain acts. And the threats that it offers may have a more indirect role in shaping people's decisions, encouraging them not to act wrongly."[393] However, Tadros does not believe that moral wrongness is a necessary condition for the permissibility of criminalising a conduct. One example he uses to show this is inchoate crimes and uses the example of possessing firearms.[394] In summary, his idea is that there are cases of possession of firearms where it is not plausible to say that an agent has committed a moral wrong – they only possess a firearm for recreational purposes and do not intend to use it for any kind of wrongdoing, for example. Even in such cases, Tadros argues, the state may be justified in criminalising possession of firearms for separate reasons, such as protecting citizens from homicide or improving overall security.[395] Instead, then, the negative role that moral wrongness is to play in a criminalisation deliberation process is as a way to determine *proportionality* for the punishment that criminalisation warrants. As he puts it, "conduct should normally be criminalized only if the punishment that criminalization warrants is proportionate to the crimes that have been committed."[396] In this sense, a decision-maker should determine whether a conduct is morally wrongful in order to then determine whether the seriousness of that wrongness can defeat other reasons that might count against criminalisation, and in doing so can also determine whether the punishment that criminalisation leads to would be proportional.

[392] ibid 159–160.
[393] ibid 164.
[394] ibid 17.
[395] ibid 333–334.
[396] ibid 170.

4.2.5.a. Applying the questions to Tadros

Since Tadros is not presenting his criminalisation theory in the book as a neat set of principles, as in some of the examples above, it is slightly more difficult to evaluate the effects of metaethical stances on his views of criminalisation. However, Tadros does offer a structure for what a potential deliberation process for criminalising *based on* moral wrongness *should* look like, which one can apply the metaethical questions to. There is also the fact that Tadros does seem to hold some particular views on the potential roles that moral wrongfulness can play in criminalisation theory, and also has a specific view on what gets to count as moral wrongness as a distinctive kind of wrongness, which are both worth exploring from a metaethical perspective.

Let us begin, though, by looking back at the reconstructed deliberation process that Tadros states. In it, the first few stages are the most relevant for the metaethical discussion I am proposing, particularly because they are the steps by which moral wrongfulness is identified and explained. The first step is, precisely, identifying *why* the target conduct is wrong. To do this, Tadros rightly tells us, we need two things: an account of the nature of wrongdoing, and an explanation for the source of wrongdoing for the target conduct specifically. In other words, we need to know what wrongdoing *is* and we need to know why Φing is an *instance* of that kind wrongdoing. Because the wrongness being sought after is of a moral kind, and because Tadros believes that moral wrongness is a distinct kind of wrongness, a decision-maker will have to determine whether the conduct in question includes facts that give us a decisive reason against carrying out the conduct. In this sense, Tadros is recognising the importance of, in the terms I have put it thus far, answering

the conceptual question for moral wrongness. When giving an answer to this question, Tadros starts by acknowledging that part of the distinctiveness for moral wrongness is that it provides a decisive reason against the conduct. However, he is clear that by 'decisive' he does not mean 'categorical' because, as he rightly points out, there are other normative contexts that can provide categorical reasons (i.e. non-hypothetical reasons) such as rationality.[397] What he does say is that, in this kind of wrongness being 'decisive', "moral wrongness dominates other kinds of wrongness in practical deliberation."[398]

Not only that, but this decisive reason must also compel us in a particular way – have a specifically moral normativity – to follow that decisive reason. I speak of 'facts' that give a decisive reason (in this context, a moral reason), because Tadros thinks that the way in which wrongness gets determined is precisely by having clarity as to what kinds of facts – and, therefore, which judgments about those facts – make it the case that an act is morally wrongful. As he puts it, "any view of wrongness relies on distinctions between facts that make acts wrong and facts that do not. […] Any theory of wrongness, […] relies on *independent judgments about moral salience* of certain facts."[399] It is precisely these independent judgments about moral salience of facts which are an important aspect of answering the conceptual question, and to which Tadros does provide us with a kind of answer: certain facts make it the case that there are decisive and especially compelling reasons against Φing. These facts are what make the wrongness involved of a moral kind, and the judgments we make about those facts are the first step that Tadros requires of decision-makers in the deliberation stage.

[397] ibid 16.
[398] ibid.
[399] ibid 24. Emphasis added.

So far so good. Difficulties start to arise, however, when we continue asking relevant metaethical questions. For example, what are the *kinds* of facts that have this particular moral salience? And, importantly, how can they be found by the decision-maker? I will look at this latter question further below, when dealing with the epistemic question. The first question, however, is trying to determine a potential answer to the ontological question, and Tadros is not explicit in his work regarding his metaethical commitments in order to answer it. So, trying to reconstruct Tadros's answer to the ontological question can be challenging. However, we do get some clues from how he views morality and how he presents some of the moral principles he uses, that can give us some idea of what metaethical commitments are being assumed. For example, in his discussions on punishment and wrongdoing, Tadros states that wrongdoers "may reject the ends that we serve by punishing them, but their rejection is not decisive where they have duties to serve these ends."[400] He also states that the behaviour undertaken by people when they plan their lives and set themselves aims "is appropriate only if the person's aims and plans are pursued within the constraints that morality sets for her" and that "a person who *values correctly* recognizes the values that constrain her from bringing about outcomes that she values."[401]

These views suggest that Tadros, at the very least, is assuming some form of mind-independence regarding moral facts. The fact that the ends pursued by 'us' in punishing wrongdoers are not altered or affected by the rejection of the wrongdoer, suggests that those ends are, at least, independent of the mind of the wrongdoer. Pairing this with the fact that Tadros states the existence of duties of wrongdoers to *serve* those ends, that are not altered by

[400] ibid 48.
[401] ibid 50. Emphasis added.

the will of the wrongdoer, suggests that those duties are, again as a minimum, independent of the mind of the wrongdoer. Now, perhaps one might think that this independence is only true as an exception regarding wrongdoers, but those moral facts about the ends of punishment might still be mind-dependent regarding 'us' as a society, in some form of inter-subjective making it the case that those facts exist because we value those ends. However, this cannot be Tadros's view because he also believes that part of the distinctiveness of moral wrongness is that it is true independent of someone valuing or disvaluing the wrongful conduct. He exemplifies this in his *Sinking Ship* thought experiment,[402] where the Captain of a sinking ship draws lots to save the children on the ship without rigging the lottery, even though his child is on the ship. His child is still saved, but it turns out that this is because the First Officer wrongfully rigged the lottery to save the Captain's child and, if the First officer had not done this, the Captain's child would have died. Tadros uses this example to suggest that the Captain can, without contradicting himself, be grateful that his child is saved and that the First Officer wrongfully intervened, and also criticise the First Officer for acting wrongly. As Tadros puts it, "he can rightly claim that he would not have acted in the same way, as demonstrated by the fact that he did not rig the lottery, and he can criticize the First Officer as a result. This does not conflict with his being glad that the action was performed."[403]

Regardless of whether one shares this intuition or not, the point is that Tadros can make this claim because he accepts that the *facts* that make the rigging of the lottery morally wrong are not changed by the Captain's valuing of his child being saved. In this sense, those facts are mind-independent even from the party that would value the ends pursued by a conduct that can be

[402] ibid 29.
[403] ibid.

seen as morally wrongful. It doesn't seem that this fact would change for Tadros if we all valued – as probably many of us would – the fact that the child is saved; rigging the lottery would still be morally wrong. And Tadros also states that this is not just a way of showing how consequentialism can be challenged (as the normative ethics theory that would propose the valuing of the outcome as the first-order judgement of moral wrongness), but also as a way to explain the "psychological conflict that morality seems to give rise to",[404] which is an issue relevant to metaethics. Specifically, it is the issue of the effect of a mental state (a compulsion to prefer the outcome of saving the Captain's child) on the existence of a moral fact (the wrongness of rigging the lottery). Tadros argues that "whilst this psychological pull is powerful he is constrained from acting on it" and that "[the Captain] recognizes this by conducting the lottery fairly."[405] In presenting this issue thusly, it is fair to say that Tadros's views are coherent with a kind of moral realism as mind-independence answer to the ontological question, even if he does not explicitly endorse this view.

Finally, let us look at the epistemic question. There is not much discussion of moral epistemology within Tadros's book, but he does dedicate some words to it regarding political liberalism as the stage under which much of criminalisation theory is developed.[406] Tadros discusses political liberalism as relevant to criminalisation because, as he puts it, "the general idea of political liberalism is that, given widespread reasonable moral disagreement in modern liberal societies, there is a reason to ensure that state institutions, laws, and acts do not depend for their justification on controversial moral views."[407] Thus, if moral wrongness is to play a role in the criminal law, political

[404] ibid.
[405] ibid.
[406] ibid 8.
[407] ibid 136.

liberalism can serve an important role to help us "understand why the deep connection between wrongs and crimes survives the disagreements that inevitably, and quite reasonably, arise in liberal states."[408]

Importantly, Tadros is sceptical of the general criticism that one can make from an epistemic perspective of moral knowledge, within the context of political liberalism. This is because, as he puts it, "epistemic considerations […] are not central to political liberalism, as it is best understood."[409] He supports this idea by saying that "our confidence in at least some moral ideas is not very high, and that confidence can be in part eroded by the fact that others, who seem reasonable, have a different view."[410] But, Tadros says, this is irrelevant to the context of political liberalism because the same can be said about the acceptance and appropriateness of political liberalism itself – there is disagreement about endorsing political liberalism among epistemically reasonable people.[411] That being the case, Tadros feels justified in stating that "I will often consider the implications that certain political and legal norms have for those with true moral views and for those with false moral views, as though we know which view is right", and that this will avoid "being distracted by questions of uncertainty that are irrelevant to the best case scenario for political liberalism."[412] We also get some additional clues to his assumed moral realism, particularly regarding the truth-aptness of moral claims, because he is sympathetic to the idea that there is no real problem with imposing a moral view, which is a common criticism based on liberalism, as long as that view is true. As he puts it, "the objection that one has acted in a partisan way if one imposes a moral view on others hardly seems powerful when truth is on one's side—it is the truth rather than the fact that the truth is

[408] ibid.
[409] ibid 143.
[410] ibid.
[411] ibid.
[412] ibid.

'on one's side' that justifies the decision."[413] This also has effects for the epistemic question, specifically regarding justifying a belief in the moral wrongness of a conduct because, as Tadros puts it and agreeing generally with Enoch,[414] "when one person imposes the correct moral view on another, this imposition does not amount simply to an imposition of one person's beliefs on another in any morally important sense. It is an imposition of the right view on the other person, and must be assessed as such. Adding that it is also the person's belief is simply a distraction."[415] Thus, we can have a clearer sense that Tadros, along with the moral realist, both thinks that moral wrongness is truth-apt and that, therefore, we can gain moral knowledge.

It also seems that Tadros is at least sympathetic to an epistemic methodology of moral intuitionism. He uses moral intuition as a way to exemplify how direct epistemic reflection on moral views is not always necessary in order to hold a moral view, and states that "we have powerful moral intuitions without experience. And yet we often disagree about these scenarios, and about the intuitive plausibility of principles themselves."[416] He also is comfortable in assessing things like moral responsibility by using moral intuitions as the first point of reference for reflection,[417] and has also appealed to intuition as a way to evaluate different understandings of his hypothetical examples in his response to criticisms to his work.[418]

With this reconstruction in mind, let us review what the deliberation stage

[413] ibid 139.
[414] David Enoch, 'The Disorder of Public Reason' (2013) 124 Ethics 141.
[415] Tadros, *Wrongs and Crimes* (n 1) 140.
[416] ibid 143.
[417] See the discussion found in ibid 75–79 on Pereboom's arguments on responsibility in light of determinism, found in Derk Pereboom, *Free Will, Agency, and Meaning in Life* (Oxford University Press 2014).
[418] See Victor Tadros, 'Responses to Wrongs and Crimes' (2019) 13 Criminal Law and Philosophy 455.

(at least for 'serious crimes') looks like for Tadros. The first step the decision-maker must take is to identify a conduct as a wrong, and in doing so explain why that conduct is wrong. To do the latter, they need two decision-making tools: a definition of the nature of wrongdoing, and an explanation as to what is the source of that wrongdoing in the conduct in question. If this wrongness is of a moral kind, then that, from Tadros's answer to the conceptual question, entails two things about the nature of that wrongdoing. First, that there are morally salient facts about the conduct that give potential agents decisive reasons against carrying out the conduct, and that the presence of those facts normatively (and not psychologically) compels the agent to refrain from the conduct. The relevant morally salient facts, in turn, are ontologically mind-independent and desire-independent facts, which the decision-maker can make true or false judgements about. Finally, the decision-maker can identify those facts, and judge them accordingly, based on their intuitive findings in first instance, while then subjecting those intuitions to closer reflective scrutiny. He would argue, however, that the possibility of moral uncertainty, at least in the context of political liberalism, is not relevant to how the deliberation process works.

4.3. Metaethical differences matter to the Deliberative Stage

As we have seen throughout this chapter, though each theorist has seemingly been using the same conceptual tool of moral wrongfulness for their theorising on criminalisation, the reconstruction of their usages has shown that there are different metaethical characteristics being assumed for that moral wrongfulness. It is fair to say that for most of the theorists discussed in this chapter, there seems to be an underlying assumption of a broadly

moral realist metaethics. For some this is more explicit, such as in Moore's theory; for others, moral realism is more of a given and is simply assumed in their talk of moral wrongness, which requires a construction of potential answers to the metaethical questions posed, such as with Tadros's theory.

However, even within their assumed agreement of moral realism, there are some important differences in how each theorist's constructed answers understand the metaethical content of moral realism. For example, though both Duff and Moore use moral wrongness as a positive reason for criminalisation, Moore thinks that said reason is found in the natural properties of the conduct in question, and that said wrongness is a very weighty reason in the deliberation process. Duff, in contrast, thinks that moral wrongness only acts as one, *pro tanto* reason for criminalisation if it is the right kind of wrongness (i.e. a public wrong), and said distinctiveness is not found in the natural properties of the conduct itself, but rather in the rationally reconstructed values that are constitutive of the civil order. Similarly, for those who only see moral wrongness as a negative constraint on criminalisation, such as Simester & Von Hirsch and Tadros, the appropriate way of finding the relevant wrongness will be different. For Simester and Von Hirsch, the decision-maker is looking, through rational deliberation, for the all-things-considered reason that makes it the case that the conduct is morally wrong; for Tadros, the decision-maker is supposed to find a decisive reason that rules out the conduct being criminalised as a valid choice of action, which they can initially do by observing their intuitive reactions to the conduct in question in order to find that distinctively moral constraining experiencing of the conduct being criminalised.

My purpose in finding these metaethically-relevant differences between theorists has been to show that not addressing the metaethics involved in

developing normative criminalisation theories leaves important gaps in our understanding of the deliberation process that we expect legislators to engage in. The differences in metaethical commitments exist and they are not being addressed by most criminalisation theorists. Not doing so leaves important theoretical gaps, that are worth addressing with the tools provided by metaethics. A second gap exists because these metaethical differences can lead to different understandings of how the criminalisation theories are supposed to be *applied*. The criminalisation theories being discussed in this thesis are presumably meant to be capable of being *used* by criminalisation decision-makers, who would go on to attempt to apply the deliberation processes developed by the theories. By not addressing the metaethical commitments that inform the theories in question, theorists are leaving another practical, *methodological* gap that decision-makers will necessarily have to fill in themselves. Additionally, other criminalisation theorists will have their own views on metaethics (sometimes unconsciously or as mere assumptions) and leaving this theoretical gap can also affect their understanding and evaluation of criminalisation theories, as I will explore further below.

Perhaps the clearest way in which the importance of the metaethical differences between these theorists can be shown – and, potentially, between theorists and decision-makers intending to use these theories in their deliberations – is by looking at an example of a candidate conduct for criminalisation and seeing how different metaethical commitments can change our understanding of whether that conduct is 'wrong' in the relevant sense for criminalisation. Imagine that we are faced with the decision on whether to criminalise a conduct that people have strong moral opinions about: cannibalism, meat-eating, incest or the consumption of heroin, to name a few possible candidates. One of the first questions that we will need

to ask ourselves as decision-makers, if we are trying to apply the deliberations proposed by these theorists, is whether the conduct in question is morally wrong. There are two first-order answers that we will be able to give: yes or no. These answers, however, do not come from a vacuum and are not, we will assume, arbitrary assertions. That being the case, we can reasonably ask a few further questions about the first-order answer we just gave. The first question we might have is about reasons why the conduct is morally wrong. Here, we might still provide ourselves with a first-order answer from normative ethics. Put crudely, one of us might say "no, because I am a consequentialist and I think that there are no net negative consequences caused by the conduct", and the other might say "yes, because I am a deontologist and I think that there is an undefeated duty to not perform the conduct". But, importantly, there is a broader sense for that "why" question, which has to do with *how we came to the conclusion* that the conduct is or is not wrong. That is, we might be asking questions more along the lines of "what do you mean by 'wrong' here?" and "how do you know that this conduct is wrong?".

The most important reason to provide these kinds of answers in criminalisation theory is because doing so will provide the basis for a clear epistemic method for the deliberation process. If we expect decision-makers to be looking for moral wrongness in their deliberations, then they need to know two crucial things: (a) what it is exactly that they are looking for, and (b) how they are going to look for it. Not providing guidance on this matter in our theorising – or at least to make our expectations clear – can lead to decision-makers filling in the gaps themselves, and that can in turn lead to very different applications of criminalisation theories in practice.

To see this, here is an example. Let us assume that there are two potential

decision-makers trying to deliberate on whether to criminalise incest in their jurisdiction, and they have at their disposal the theories explored in this chapter as potential frameworks for deliberation. Further, let us assume that one of them, call them Ash, holds views on morality that are consistent with robust moral realism; Ash thinks that there exists a matter-of-fact about moral wrongness, and that those facts are independent of their own thoughts or desires about that wrongness. Further, let us assume that Ash is also a naturalist realist, and therefore Ash believes that they need to determine which natural properties of a conduct can correctly track moral wrongness. Now, let us assume that the second decision-maker, call them Bree, also holds the view that moral wrongness is a matter-of-fact, but does not think that it is a mind-independent fact and, instead, believes something along the lines of constructivism; that from holding a particular practical standpoint, moral facts can be constructed when they are coherent with holding that standpoint. Therefore, Bree is more inclined to use a coherentist reflective equilibrium in order to determine what facts can be constructed based on the practical standpoint that they are occupying as criminalisation decision-makers, along with the other norms and values that they might bring into the deliberation based on other roles they play within the jurisdiction, such as being a member of the polity, being a parent, being a legislator, and so on. Finally, let us assume that they both agree to begin their deliberation based on Moore's deliberation model presented in "Placing Blame". So, they both accept that if they can determine that incest is morally wrongful, they have a weighty reason in favour of criminalising it.

Ash might think something like the following: "I believe that incest has the property 'moral wrongness' because I have evidence that incestuous relationships lead to genetic malfunctions through reproduction that I can observe, and I can ground moral wrongness on those natural facts that give

me a sound causal explanation to taking the conduct as morally wrongful. Therefore, I have a positive reason to criminalise incest." One potential response that Bree could have would go something like the following: "I don't believe that there are any moral facts I can reasonably construct for the property of moral wrongness for incest, because I do not have any evidence for such a moral fact to be constructed from the practical standpoint I hold, and therefore I have weightier reasons in favour of the moral permissibility of incest. Therefore, I have no positive reason to criminalise incest." These are obviously simplified versions of their thought-process, but they are enough for my purposes here, for even in their simplified form we can already see a potential issue: both decision-makers are using the same deliberation framework but reaching different conclusions. And those different conclusions are *not* based on first-order disagreements on the moral character of incest. Instead, they are disagreeing on how that moral wrongness is *determined* in the first place, because they hold different views as to what *counts* as the property moral wrongness that they are seeking. They also disagree on which epistemic method they should use to make up their minds about that moral property.

Interestingly, if Ash and Bree are trying to follow Moore's deliberation process, Ash has an additional point to make in the discussion, because they hold the same metaethics as Moore does: naturalist moral realism. So, Ash can make an additional criticism to Bree regarding their disagreement: that Bree is not following the metaethical commitments that ground Moore's deliberation process. And, as they have both used the same deliberation process but reached different conclusions, Ash can at least argue that their conclusion is more consistent with how Moore views the moral deliberation process and, in view of that, that their criminalisation deliberations are more in-line with how the theory is supposed to be applied.

In conclusion, this chapter has shown that addressing the metaethical commitments involved in normative theorising of criminalisation is both important and useful for how criminalisation theories can come to be used in practice. It has shown that some of the most prominent criminalisation theorists have different answers to the metaethical questions posed by this thesis, and that those different answers shape the deliberation process for criminalisation in important ways. Thus, I have argued that it is important for future criminalisation theories to be both aware of metaethics in their normative theorising, as well as being more explicit with what metaethics are being assumed in the theorising of criminalisation. In particular, I have shown that having this metaethical clarity has two main payoffs: a better understanding of the criminalisation theory in order to evaluate it more fully from a theoretical perspective, and a better understanding of how the theories can end up being used in practice by potential criminalisation decision-makers.

CHAPTER 5 – THE POLITICAL TURN THEORIES OF CRIMINALISATION: AN ALTERNATIVE TO LEGAL MORALISM?

I wish to now pay some closer attention to a different trend in criminalisation theorising, that has come about as a sort of response to the recent trend of legal moralist theorising. These responses have focused primarily on shifting the perspective from which to observe the criminal law from morality to politics, although this can be done by different means and with different focuses. As Lacey and Pickard have put it, "in recent years there has been what is often referred to as a 'political' or 'public law' turn in criminal law theory, developing a wide range of purposive, regulatory interpretations of criminal justice."[419] This entails, in a nutshell, making an argument based on a particular view of what the criminal law is for, which is contrary to the legal moralist account of seeing the criminal law's function as one of moral blame or censure. Lacey and Pickard state it thusly: "the idea of standing to blame as a distinct condition on holding to account cannot simply be assumed by analogy with the juridical model of interpersonal moral blame."[420] This 'political turn' can be specifically observed in the context of criminalisation theorising, though it does not embody a defined set of theses about this topic. Instead, it seems to be a *kind* of response that a theorist might offer to legal moralist criminalisation theories. This response is to deny the legal moralist's claim about the relevance of *moral* wrongdoing to criminalisation, and propose as an alternative a 'political' account of normative facts that are relevant to criminalisation. There are various different ways in which these relevant normative facts can be defined and identified,

[419] Nicola Lacey and Hanna Pickard, 'Why Standing to Blame May Be Lost but Authority to Hold Accountable Retained: Criminal Law as a Regulative Public Institution' (2021) 104 The Monist 265, 275.
[420] ibid.

but I will limit the scope of my analysis here to some of the more influential 'political turn' literature about criminalisation.

The purpose of this chapter is to apply the metaethical lens I have been using for theories that use moral wrongness, and turn that lens towards some examples of these political turn theories. My objective is to provide further tools from metaethics to better understand the project of the political turn theorist, and in particular to better understand how, if at all, a political turn theory can be different to legal moralism and theories that emphasise the use of moral wrongness. I do this by analysing some recent contributions to the political turn literature regarding the criminal law, and explain different ways in which using the tools of metaethics could benefit the theorising of the analysed authors, as well as show how *not* discussing the metaethical implications of a 'political turn' theory of the criminal law leaves theoretical gaps that are important to acknowledge.

I then claim that there are two ways in which we can make sense of the contrast being made between the 'moral' and the 'political' by the political turn theorist, which I will call Option 1 and Option 2. Option 1 is to claim that the 'political' is a *subset* of norms and reasons within the domain of morality. In this sense, the political turn theorist will accept that the wrongfulness that is either being expressed through criminalising, or that a decision-maker should consider in their deliberation process, is of a moral kind, but that it is of a specific part of morality that is distinct, which we can call 'political morality'. Option 2, in turn, is a stronger claim than option 1, whereby the political turn theorist will claim that the 'political' is a *distinct normative domain*, one which is separate and different from morality. There may be connections between the 'moral' and the 'political', but Option 2 theorists will

want to claim that those connections are not a necessary conceptual feature for the political domain, and that the two domains are conceptually distinct.

An upshot of having this clarity is that it allows for a suggestion, which is that if one wishes to be a 'political turn' theorist *and* provide an alternative to legal moralism, one needs to choose between bullets to bite. One choice is to bite the bullet on accepting that the moral and the political are not as separate as the distinction may make them seem, thereby reducing the distinctiveness of political turn theories and instead presenting the theory as an alternative to 'traditional' legal moralism, thereby becoming a distinct variety of legal moralism (specifically linked to 'political morality'). The other choice is to bite the bullet on morality and politics being different normative domains, thereby offering a fully alternative normative theory to legal moralism, but running into the challenges posed by successfully distinguishing the political normative domain from the moral one.

5.1. Vincent Chiao and Fully Political Standards

I will begin by analysing one of the more recent and extensive efforts within this line of thought. Vincent Chiao, in his book "Criminal Law in the Age of the Administrative State", sets out to present an alternative to legal moralism as an understanding of criminalisation. In particular, Chiao contrasts his view to retributivist legal moralism, which he takes to focus on, as he puts it, the "vindication of private rights – the rights people would have in the state of nature."[421] This description of legal moralism is in line with Farmer's view that "most contemporary theories understand the scope of the criminal law as being properly determined by some combination of the rights or interests which are to be protected and the identification of harms or wrongs […] to

[421] Chiao (n 269) 27.

those rights or interests."[422] Chiao, in contrast, argues that the criminal law has a much more important function than the protection of individual rights, and instead posits what he sees as "the central moral concern with modern criminal law: its legitimation as a *political* institution."[423] In this sense, Chiao sees the main function of the criminal law as "stabilising social cooperation with public institutions by providing assurance that those who cooperate will not be met with defection and exploitation."[424]

To defend this view, he begins by making the above claim about what the criminal law is actually for, and situates it within the political context of what he calls the "modern administrative/welfare states".[425] Though he is willing to accept the possibility that a legitimate function of the state could be ensuring that (moral) wrongdoers receive deserved punishment and censure, he states that "what I am at pains to reject, however, is the thought that a political morality of this kind could be sufficiently motivated simply by attending to the moral value of retribution", and adds that "there are lots of ways of responding to wrongdoing, and the choice among those responses is, or so I shall argue, a substantive political decision for the citizens and officials of a given polity."[426] From there, he builds on an account of "criminal law as public law" to argue that "the criminal law is first and foremost an institution that contributes to making social cooperation possible, regardless of whether it is or is not also an institution devoted to the pursuit of retributive justice."[427] In this sense, and relevant to criminalisation specifically, Chiao makes great efforts to resist *justifying* the contents of the criminal law simply by appealing

[422] Lindsay Farmer, *Making the Modern Criminal Law - Criminalization and Civil Order* (Oxford University Publlishing 2016) 16.
[423] Chiao (n 269) 28.. Emphasis in the original.
[424] ibid 51.
[425] ibid 36.
[426] ibid 31.
[427] ibid 36.

to the fact that retributive justice is supposedly being achieved by punishing wrongdoers, and instead wants us to focus on the political justification that criminal law requires as a public institution. Thus, Chiao proposes that the criminal law requires what he calls a "fully political standard of justification"[428] in order to properly justify the contents of the criminal law and, importantly, criminalisation.

His main argument regarding how the criminal law is justified by a "fully political standard" is that "if the primary function of the criminal law is to support social cooperation with public institutions, then a moral evaluation of the criminal law should be drawn from an account of our reasons for supporting public institutions in the first place, rather than directly from the norms of interpersonal morality."[429] Note, however, that Chiao is still interested in how the criminal law should be *morally* evaluated. He is not suggesting that the norms by which we evaluate criminal law practices – criminalisation being one of those practices – are non-moral norms. Instead, he seems to be suggesting that there is a distinct set of moral norms whose *target* is politics and political life, and that this is different from "inter-personal morality" norms. He differentiates the two by saying that the former is constituted by "moral principles designed to evaluate the operation of complex and law-governed institutions that claim authority over large and diverse populations", and contrasts this with the latter inter-personal morality as being used to evaluate "face-to-face interactions between private individuals acting on their own consciences."[430]

Before looking at the contents of this 'fully political standard' for justification, and its application to the context of criminalisation, I first want to

[428] ibid 51.
[429] ibid.
[430] ibid 64.

apply the metaethical lens to this distinction that Chiao is making between 'political' and 'inter-personal' morality. To see if this distinction makes sense, we can apply some of the metaethical questions I have posed in earlier chapters and see what metaethics needs to be assumed for the distinction to work. It seems that Chiao is offering two different answers to the conceptual question on what counts as 'morality'. On the one hand, we have political morality, which is constituted by the standards under which we can evaluate political institutions. On the other hand, we have inter-personal morality, which is constituted by the standards under which we can evaluate individual agents. These would be the targets for each moral domain, though Chiao does not give us much more than that as to how these domains are conceptually distinct. This can lead us to quite a few conceptual gaps, which again can be addressed by appealing to ideas in metaethics. An initial gap, for example, has to do with what exactly is being evaluated in each domain. That is, when evaluating individual agents we can look to, for example, their *actions* or their *character* and make moral judgments; when evaluating political institutions, what should we be looking at? Possible analogues might be institutional *practices* and institutional *design*, but more needs to be said as to what constitutes these analogues, and how those aspects of political institutions provide reasons for or against criminalising conduct.

A more important question to address, however, has to do with the normativity of each domain. It is fair to say that Chiao would agree that when evaluating something like the criminalisation of conduct, the moral evaluation of the *act* of criminalising from a standard of political morality is more authoritative than the moral evaluation of the *conduct* that is being criminalised. This is because he advocates for the *justification* of criminalisation from a 'fully political standard' for that (moral) justification, so presumably a reason that can be obtained from that political standard will be

more authoritatively normative than a reason obtained from the inter-personal standard. Again, for this view to make sense, one needs to accept the possibility that there can be a distinct normativity for political morality, as understood here by Chiao, that in regards to the criminal law will be deemed more authoritative than the normativity from inter-personal morality.

Chiao presents his 'fully political standard' as follows:

> "The criminal law is worth supporting if and only if:
> a. The institutions whose rules it enforces are worth supporting, and
> b. Its use in a particular context would be consistent with the principles that make those institutions worth supporting in the first place."[431]

Several issues relevant to metaethics can be extracted from this statement of the standard, particularly to the assumptions Chiao is making on the authoritativeness of this "justification". Notice that Chiao uses the terms "worth supporting" rather than "is justified". This language suggests that the authority of the justification for the criminal law under this standard is one more akin to partisanship than, in the language used in earlier chapters, inescapability. Indeed, this way of understanding what (a) and (b) are providing is consistent with the idea that the deliberative weight of a moral fact, such as an action being "worth supporting", is a matter of *choice* of the decision-maker, rather than an imposition on their deliberations from something like rationality itself.[432] Again, this is a metaethical assumption

[431] ibid 53.
[432] Similar to the questioning of moral supremacy questioned by Tiffany in his defence of normative pluralism in Evan Tiffany, 'Deflationary Normative Pluralism' (2007) 33 Canadian Journal of Philosophy Supplementary Volume 231.

being made which, regardless of whether one agrees with it or not, would benefit from being made explicit as part of the theory being offered by Chiao.

Another important metaethical issue has to do with where the norms and reasons that feature in Chiao's conditions (a) and (b) *come from*. For him, the terms under which the justification of the criminal law is carried out "are consistent with how public institutions generally are justified" and "the same normative standard applies in both (a) and (b). Our reasons for deeming a public institution as worthy of support in (a) apply, in (b), to determining whether such an institution would itself support the use of the criminal law as a means of rule-enforcement."[433] So, for the justification being offered by Chiao's standard to be 'fully political', there are at least three normatively relevant elements of the standard that *need not* be linked, somehow, with what he calls inter-personal morality: the judgment of 'worth supporting' for the institutions in (a); the judgment of consistency in (b); and the principles that make institutions worth supporting in (b). And, again, for this to be the case, important metaethical assumptions need to be made. The most important one I will mention here is the idea that the normative *facts* (such as the norms or reasons) that form each standard can be intelligibly seen as *distinct* from each other. That is, there will be a set of 'political-moral' facts that form the fully political standard, and there will be a *different* set of 'inter-personal-moral' facts that form a moral standard. Here is, precisely, the first approximation that Chiao could be making to show a difference between the political and the moral: it is not just a matter of scale of the relations covered by these domains, but rather from the *source* of the norms and reasons being used to justify the practices of the criminal law. These different sources are, supposedly, self-contained sets of moral facts that are relevant to different kinds of relationships amongst humans and, importantly, among institutions.

[433] Chiao (n 269) 54.

It is clear that Chiao believes in this separate nature of these domains, since as he puts it:

> "the norms of interpersonal morality are not the only way to ground a normative theory of criminalization. One might, for instance, start with a conception of justice in public institutions. Adopting such a perspective would naturally lead one to consider when just institutions would approve the use of the criminal sanction to enforce their laws and policies. This would provide a basis for critical evaluation of the criminal law."[434]

Again, the point of showing the above is less about being critical or supportive of the conceptual understanding that Chiao is using here. Rather, my point is to show that *none of this* is discussed in his theorising – it is all simply assumed. This is problematic for two reasons. First, because this is by no means an uncontested assumption to make.[435] And, second, as we have come to see throughout this thesis, these assumptions matter for our proper understanding and evaluation of criminalisation theories. One reason it is important to have this clarity, for example, on the authoritativeness of reasons from political standards, as opposed to inter-personal, has to do with considering the possibility for conflict between the two standards, if that conflict is accepted as intelligible. We first need to ask ourselves whether it even makes sense to think that, for example, an action might be considered permissible under one standard and wrong under the other: can you do something politically right and inter-personally wrong? More specifically, can you deem an action inter-personally permissible or right, yet politically wrong, and vice versa?

[434] ibid 159.

[435] For a recent criticism of the view in favour of a distinct political normativity, see Jonathan Leader Maynard and Alex Worsnip, 'Is There a Distinctively Political Normativity?' (2018) 128 Ethics 756.

One possible answer Chiao might give to this issue is that, at least for the context of criminalisation, it is irrelevant to the decision to criminalise a conduct, since as he puts it "the moral limits of the criminal law should be found not in *what* is being regulated, but in *how*: not in the moral quality of some act considered in isolation, but in the principles governing when it is appropriate to rely on ex post sanctions to ensure robust and ongoing social cooperation."[436] That is, the target of the decision to criminalise is no longer the conduct itself, but rather whether said decision is consistent with the principles that make political institutions worthy of support. But this answer is too quick because, as we shall see below when looking at Chiao's deliberation process, the way in which the appropriateness of the measures applied to ensure social cooperation is determined is not independent of the normative status of the conduct being criminalised. In fact, Chiao recognises this himself, regarding what he calls the "subject-matter constraints" that focus on the normative status of the conduct being criminalised, and states that "perhaps some type of subject matter constraints might be defended from within a public law perspective as part of a more general view of justice."[437] So, at the very least, the normative status of a conduct is in some way connected to the evaluation of the actions of political institutions, which means that the possibility of there being a conflict between the two kinds of normative standards cannot be dismissed, at least in some cases, as irrelevant. Thus, teasing out the distinction between the two standards, and thereby the two moral domains, becomes an important task for both Chiao and any theorist that seeks to use a similar distinction.

[436] Chiao (n 269) 150. Emphasis in the original.
[437] ibid 161.

As for other metaethical issues that have been explored in this thesis, as one might suspect, Chiao does not provide any potential answers to the ontological or epistemic questions posed earlier. This leaves unanswered a very important question about the fully political standard for justifying criminalisation: how is the criminalisation decision-maker supposed to *find* that standard? Let us look at what Chiao proposes as the deliberation process a decision-maker would need to go through when following his fully political standard, and see if we can find any clues. Chiao describes the deliberation as follows:

> "Under a fully political standard, deciding whether it is appropriate to criminalize X requires determining whether doing so stabilizes cooperation with worthwhile institutions, that being the basic point of using the criminal law to enforce public rules; and, if so, further determining whether the use of criminal sanctions to achieve that end would be consistent with our reasons for considering that institution to be worth supporting in the first place. This latter calculation would necessarily include consideration of the potential for intrusive surveillance, police harassment, excessively onerous compliance requirements, and the other ancillary effects of criminally enforcing a rule."[438]

Similar metaethical issues that I have discussed above become salient for this type of deliberative process. The more obvious ones are, again, regarding the conceptual question of what gets to count as the relevant 'morality' here, particularly when it comes to determining the 'worth' of political institutions. Here is one way of framing the problem. It seems that Chiao's view of political institutions 'worth supporting' are, precisely, political institutions that do not wrong their citizens – they stabilise coordination by

[438] ibid 160.

not doing practices that, in some way, wrong the members of a polity. If this is the case, then what is the conceptual nature of that wrong? Is it different to 'inter-personal' moral wrongness and, if so, in what way? As explored above, the difference is probably going to be argued to lie in the source of the norms and reasons under which a judgment of 'worthy of support' is reached, but there is nothing obvious about the claim that said norms and reasons are not simply the same ones as those we would use to judge moral wrongdoing at an inter-personal level – it needs to be argued and explained. This is particularly salient when Chiao speaks of how one needs to determine whether the use of criminal sanctions to achieve cooperation is 'consistent' with the reasons by which we judge the political institutions as worthy of support: this is simply another way of saying that it would be *wrong* to impose a criminal sanction that is inconsistent with those normative reasons. Chiao might want to resist this label and say that it would be *unjust* to impose that criminal sanction,[439] but in the absence of a substantive metaethical account of what is the conceptual difference, in this context, between 'politically-morally wrong' and 'politically-morally unjust', I do not see why we cannot speak of wrongness in this context. So, when we ask ourselves why it would be wrong to do so, moral norms and reasons that maybe seemed like belonging to the 'inter-personal' side of things can start creeping into the 'fully political' moral standard, such as desert, blameworthiness and so on. And yet, Chiao is resistant to the idea that wrongness and wrongdoing, inter-personal or otherwise, has *any* role to play in a decision to criminalise a conduct. When discussing wrongfulness as a constraint on criminalisation, he states that there might be "further deontological constraints on criminalization, but wrongfulness is not one of them."[440]

[439] See his discussion of the broad sense of 'wrongful' for the wrongfulness constraints, where its use as "conduct that is wrongful only because prohibited by just public institutions" is, to his mind, simply to "restate the claim that the offense-creating law is just, except in the language of right and wrong" in ibid 172.
[440] ibid 181.

Chiao ends up presenting his proposed criminalisation principle, under which a decision to criminalise is supposed to be made, as an 'anti-deference principle' which he states as follows: "ADP+: The only reasons that would make it permissible to criminalize X are those that tend to show that criminalizing X is the best available means of promoting universal and effective access to central capability on terms acceptable to social and political equals."[441] An important element of this principle that requires some explanation is the idea of "central capabilities", which Chiao draws from a breadth of political theory literature,[442] and which he describes as "functionings that are available to a person, where a functioning is a state or activity that contributes to a person's well-being",[443] and which is 'central' because "it is required, within a given social context, to secure life without shame" and which includes things like "the ability to move about freely, to obtain a decent education, to speak freely, and to have one's voice heard on matters of public importance, to stay nourished and maintain one's health, to live and work in safe and sanitary conditions, and so forth."[444] Presumably, Chiao believes that these are all reasons that are part of the fully political standard, and therefore a part of political morality rather than inter-personal morality. However, it seems quite plausible that one could make the argument that, conceptually speaking, things like the ability to move about freely or to maintain one's health are *also* reasons to be included into an inter-personal moral standard. Surely it is the case that, if something like personal freedom or personal health were to be an actual moral fact about human life, then moral claims made about those things can be done *both* at the inter-personal as well as the political level. So, again, we require a clearer answer

[441] ibid 167.
[442] See the literature cited in fn. 30 in ibid 194.
[443] ibid.
[444] ibid 195.

to the conceptual question posed by metaethics. Since Chiao does not provide these, there is little more that can be said about the further ontological or epistemic questions. That does not mean, however, that once we *did* have more clarity as to the conceptual meaning of 'political morality' and its normativity, that we would not benefit from answering the ontological and epistemic questions.

5.2. Malcolm Thorburn and Criminal Law as Public Law

Malcolm Thorburn has had an influential role in the works of political turn scholars interested in the criminal law, including Chiao. Thorburn presents what he calls a 'public law account' of criminal justice, and explains it by stating that "my account is an attempt to justify the workings of the criminal justice system by demonstrating that they are just what is required for us to be true to a set of roles and relationships that have intrinsic values."[445] Importantly, however, "the relevant roles and relationships for criminal justice are not those we understand from ordinary morality" and instead are "the legally defined roles – such as private citizen, police officer, judge, etc – that we take up within a larger constitutional order".[446] Though he does not develop a criminalisation theory, the dichotomy he presents in his work, which inspires a large part of those who do propose criminalisation theories in the political turn, is worth analysing more closely and seeing how metaethical tools can help us better understand it.

To begin, Thorburn is not against the idea of wrongness in a broad sense as having a part in explaining the functioning of the criminal law, but he makes it clear that the wrongs involved in criminal offences, to his mind, are not

[445] Thorburn, 'Criminal Law as Public Law' (n 268) 24.
[446] ibid.

necessarily the same as, or simply an extension of, 'private' moral wrongs. For Thorburn, "criminal wrongs are all concerned with attempts to usurp either the private jurisdiction of another person or the public jurisdiction of the people as a whole"[447] And this, to his mind, is different to moral wrongdoing, and he points to existing criminal law doctrines in the Anglo-American common law world to support this view.[448] Specifically, he links this difference in the fact that, for criminal wrongs, "the law's focus is on the allocation of jurisdiction rather than on the morality of particular acts."[449] In this sense, as Thorburn puts it, the structure of criminal wrongdoing "is concerned with the protection of jurisdiction, both public and private, rather than with the identification of moral wrongs."[450] In this context, 'jurisdiction' can be read as the holding of a normative position in virtue of which an agent is legitimately entitled to determine and decide over some matter. We can see this reading in how Thorburn contrasts jurisdiction to the moral judgment one might make about a certain action, when he compares the two while speaking of civil disobedience. There he says that, in the context of civil disobedience, "whether or not we are right about the morality of the situation, it is simply not within our jurisdiction to decide such matters."[451] He also argues that this understanding of what offences do, as determining a criminal wrong based on a lack of jurisdiction, is "clearly at odds with the legal moralists' focus on the moral wrongness […] of the accused's conduct" and that his view provides the basis for "an alternative account of how we might be able to go about justifying the operations of the criminal justice system."[452]

[447] ibid 41.
[448] ibid 32–36.
[449] ibid 33–34.
[450] ibid 31.
[451] ibid 33.
[452] ibid.

Thorburn then proposes a model of liberal constitutionalism to provide a framework for the justification of the use of coercion. He tells us that, under this framework, "the state (through its officials) is not concerned with the moral rightness of our conduct *tout court*. The state's concern is with ensuring the conditions within which it is even possible for us each to make moral choices without thereby undermining our own status as the equal of those around us."[453] He then adds that "the state does so (through its officials) by patrolling the boundaries of each person's jurisdiction, using coercive force to resist any attempt to usurp the jurisdiction of others."[454] In this sense, "the use of force by state officials is justified insofar as it sets out the necessary preconditions to a life in community with others as free and equal moral agents."[455] Again, here we can see that Thorburn is at least friendly to the idea, similarly to Chiao, that there is an intelligible distinction between the political morality under which the state is providing the normative framework for public life (and does so with, at least in part, legitimised coercion) and the inter-personal morality under which we, the members of the polity and of public life, are now free and equal to act as agents.

This sympathy for that kind of metaethical distinction is even more salient when we look at Thorburn's understanding of how the state works. Thorburn's view is strongly dependent on a particular liberal view of the role of the state, specifically with regards to liberal constitutionalism as a response to Hobbes's state of nature.[456] He summarises the view by saying that "the point of the constitutional state, then, is not to accomplish some important moral task such as maximising utility or punishing moral wrongdoing. Rather, its purpose is the much more basic one of setting the framework within which

[453] ibid 42.
[454] ibid.
[455] ibid 43.
[456] See Sec.III.A in Thorburn, 'Constitutionalism and the Limits of the Criminal Law' (n 268).

we may interact with others in a way that preserves the freedom and dignity of all. The state in this model takes no role in ensuring that individuals exercise their freedom in the best way possible."[457] In this sense, the normative role of the state is, according to Thorburn, more instrumental than moral: it is supposed to be laying the normative groundwork under which further moral life can flourish among equals. Again, this 'instrumental' role is still normative, but it is of a different kind of normative domain than that of 'moral life'.

A key move that Thorburn makes to support the view that 'state-dominated practices' and 'private moral practices' are distinct is regarding how the former are a necessary condition for the latter. For Thorburn, the state and its normative context is a pre-condition for people to be able to actually have the possibility to live a moral life in the first place, specifically regarding their equality as agents. As he puts it, "in a world without [the law and the state], I simply do not have the tools at my disposal to act morally without undermining my claim to equal moral status with others."[458] As such, the state therefore acts as a way for citizens to be able to choose how to act freely, or as Thorburn says, "the state guarantees us all that we can focus on the morality of our own choices without having to worry about whether this will undermine our status as moral equals of those around us."[459]

The point of summarising Thorburn's views has not been to defend or criticise his understanding of the criminal law and criminalisation. Instead, the point has been to explain his view as clearly as possible, to now see how a metaethical lens could improve or inform his view. There are, broadly speaking, two main issues in his theorising that would benefit from more metaethical clarity. The first is similar to Chiao's issues with the distinction

[457] ibid 99.
[458] Thorburn, 'Criminal Law as Public Law' (n 268) 42.
[459] ibid.

between 'political' and 'inter-personal' morality. In not discussing how he would answer the conceptual question, Thorburn is leaving the door open to a criticism from the legal moralist along the lines of saying "we are talking about the same thing". Unless there is a substantive argument about the metaethical and conceptual differences between the two putative normative domains, or on the differences in the normativity of each sub-section of morality, then the legal moralist can simply respond by saying that what people like Thorburn and Chiao are seeing as separate normative arenas is actually just the same thing. What is more, a metaethically informed legal moralist might make a strong argument on this, by basically pointing out that the ontological nature and epistemic processes by which one identifies 'political' moral facts are *the same* as the ones used to identify 'inter-personal' moral facts, so a stronger argument is required to make the distinction intelligible.

The second issue in Thorburn's theory that would benefit from a metaethical approach has to do with the distinction he makes between 'state-dominated practices', under which criminal justice is supposed to be used, and 'private moral practice', under which people are, distinctly, called to account by one another. That is, Thorburn tries to argue that doing justice in the context of the criminal law is distinct from the practice of calling someone to account for their wrongdoing. Regardless of whether one agrees with this particular view – and I certainly disagree for the reasons I have explored in chapter 3 regarding the responsibility fact of criminalisation claims – Thorburn needs to say more about the metaethical nature of this distinction for it to stick. In particular, in view of the fact that he has not explored the issues around the conceptual question, and therefore has not made a case for the distinction between normative domains clearly, claiming that the practice of 'justice' and the practice of 'calling to account' are not the same is

slightly question begging. That is, it is perfectly plausible to argue that there might be something conceptually distinct about the kind of normative facts that ground the concept of 'justice' in a political context that are different from the normative facts that ground the concept of 'calling to account' in a 'private' or 'inter-personal' context, but that argument needs to be made explicitly, not assumed. It is also perfectly plausible to hold the view that what Thorburn is portraying as distinctive of the practice carried out by the criminal justice system – legitimate coercion – can *also* be understood as an intelligible moral practice.[460] Thus, Thorburn's understanding of the criminal law, and its distinctiveness as a political practice, could also benefit from closer attention to metaethical questions applied to his theory.

5.3. Alan Brudner and Political Theories of Criminal Law

In his book "Punishment and Freedom - A Liberal Theory of Penal Justice", Alan Brudner presents his own view on how punishment — and thereby the criminalisation that leads to it — can be justified in the context of the criminal law. Though I will not delve into the theory itself, I do wish to bring attention to a distinction Brudner makes on the *kind* of theory that he is presenting, because this distinction can be useful to see what 'political' criminalisation theories might look like, and is similar to the distinctions made by Chiao and Thorburn above.

[460] Thorburn himself cites Gardner's view of seeing punishment (a distinctive form of coercion) as something that can be viewed in the 'private' context, such as someone punishing their estranged spouse, fully intending it to be a terrible experience, which is found in the Introduction to HLA Hart, *Punishment and Responsibility - Essays in the Philosophy of Law* (Oxford University Press 2008).

Brudner states that "the best theory of the penal law's[461] general part is […] part of a political and constitutional theory concerning when state coercion of the individual is legitimate."[462] He contrasts this with an undesirable theory of criminal law that focuses on 'morality', by saying the following:

> "It is not a theory about how the moral good of punishing evil is best served; nor is it a theory about when it is appropriate to express moral condemnation of blameworthy conduct or character through the penal power of the state; nor is it an account of what it takes for punishment to be a form of moral pedagogy addressed to a socially responsible agent rather than a tool for manipulating the behaviour of a rational hedonist. To put the point succinctly: the best theory of the penal law is a theory of penal right rather than one of penal morality."[463]

He adds a caveat to his position by stating that "penal law theory is not coextensive with political theory" and, instead, is "a narrow theory about just punishment (and penalisation) rather than a broad theory of the legitimate state authority to coerce."[464] That is, though the theorising of the criminal law is more 'political' than 'moral', this does not mean that a criminal law theorist needs to explain how politics as a whole should work and how the state should use its power more generally. Instead, it is a specific application of political theory to the institution of punishment.

Importantly, however, while Brudner's account of criminal punishment is, supposedly, distinct from a moral analysis of punishment, he very openly

---

[461] Brudner uses 'penal law' to refer to the criminal law as a whole – including *mala in se* and *mala prohibita* offences – and reserves 'criminal law' to refer to the 'true crimes' of *mala in se*. See Alan Brudner, *Punishment and Freedom - A Liberal Theory of Penal Justice* (Oxford University Press 2009) 9.
[462] ibid 15.
[463] ibid 15–16.
[464] ibid 16.

accepts that it is still a normative endeavour: "Although I characterise the legal theory of criminal desert offered here as nonmoral, I do not mean by this that it is non-normative – that its subject matter is simply the fact that constitute criminal liability in some technical, value-neutral sense adapted to the managerial task of crime control."[465] So, as in the cases discussed above, it becomes important to determine *where* those values that form the normative standard for just punishment and penalisation come from, especially if they are, as Brudner puts it, 'non-moral'.

He states that his book "defends a position contrary to the dominant view for a very simple reason", which he explains by saying that "the forcible restraint of an individual's liberty by equal human beings requires a justification of an order altogether different from one suitable for expressions of condemnatory moral opinion, which the addressee is free to internalise or deflect" and that "allowing a justification for moral blame and censure to count as a justification for physical restraint cannot sit with the equal worth of agents".[466] Interestingly, Brudner also reveals at least part of his metaethical assumptions in his theorising of the criminal law, when he discusses the use of moral blame and censure as a justification for punishment. He states that "human judgments of another's inward will or character are too subjective, too plagued by uncertainty and bias, to be enforceable against an equal", and that "because such judgments are subjective and (at least partially) ignorant, the object of moral censure is always *entitled* to reject it, either in whole or in part."[467] Crucially, he contrasts this with the coercion of state punishment, and states that "the legitimacy of penal force depends precisely on the convict's *not* being entitled to reject it. And so the justification of penal liability must be a *sui generis* one geared to the requirement of non-

[465] ibid 16–17.
[466] ibid 17.
[467] ibid. Emphasis in the original.

rejectability."[468] From this, we can infer that Brudner believes moral judgments to be of a 'subjective' nature, and therefore that even if there were 'objective' facts about the moral status of agents (of which Brudner does not say much), the judgments that people make based on those facts are still 'subjective'.

As was the case with Chiao and Thorburn, Brudner is also making a normative distinction, but his is between what he calls 'just coercion' and moral blameworthiness. Since Brudner states that his theorising is non-moral, then presumably this 'just coercion', to him, is not a moral concept. It is definitely a normative concept, but it is distinct from morality, for Brudner, because it is a narrow application of political theory to state punishment, and because the justification for that punishment must be of a different 'order' of justification than the one used to justify moral condemnation, based on the fact that those who are judging the justification of punishment towards the offender are equals with that offender. However, Brudner's position becomes quite unclear once we start applying some of the metaethical questions used so far. Similarly to Chiao and Thorburn, Brudner does not explore the metaethical implications of making this distinction between just coercion and moral blameworthiness. There is a hint of what Brudner might be thinking of in this distinction, when he says that his legal theory of just punishment can be called "a moral one of a certain sort, namely, a morality of the right rather than one of the good."[469] However, this is surely not enough as a basis for the distinction, mainly because 'right' and 'good' – and its contraries 'wrong' and 'bad' – are still plausibly within the same normative domain that is being spoken of. Additionally, the kind of hard distinction that Brudner makes between 'right' and 'good' here (and of its contraries) is contestable, because

[468] ibid 18. Emphasis in the original.
[469] ibid 17.

both sets of concepts have mostly been understood as referring to reasons regarding actions, and as Finlay puts it, "that an action is 'good' just in case there is most reason for doing it, and that it 'ought' to be done just in case there is most reason to do it, are said to be platitudes".[470] But let us assume that Brudner is trying to point to an actual normative distinction between the normative facts his theory is interested in.

One way this distinction might be teased out from Brudner's work is his Kantian-inspired view of formal agency; the idea that the self is "the final end of action pursuant to its ends—that for the sake of which any action is performed."[471] This formal agency is an *a priori* capacity[472] and is 'formal' in the sense that it considers "the sole bearer of legal rights the capacity for freely choosing ends."[473] Under this view, "wrong is simply the pre-emption of, or interference with, the capacity to choose ends",[474] and perhaps it is this conceptual understanding of 'wrong' which Brudner is trying to contrast with 'bad', which is supposed to be something different. The problem, however, is that that difference is not obvious at all – why would this *a priori* capacity to freely choose ends *not* be, for example, relevant to moral blameworthiness and *only* be relevant to just coercion? How is it that pre-empting or interfering with the capacity to choose ends is a 'public' justification for punishment, but not the basis for moral censure? And, importantly, why would a judgment about pre-empting or interfering with the capacity to choose ends *not* be a 'subjective' judgment that object of that judgment – in the case of the criminal law, a defendant – is *not* always entitled to reject? The only effective way to answer these questions is to delve into the metaethical implications of making

[470] Stephen Finlay, *Confusion of Tongues: A Theory of Normative Language* (Oxford University Press 2016) 85.
[471] Brudner (n 461) 24.
[472] ibid 25.
[473] ibid 28.
[474] ibid 31.

the distinction in the first place, in order to be clear as to what *actually* (if at all) makes the normative exercises done in the domain of justification for the criminal law different from the normative exercises done in the domain of justification within morality. And again, these implications might be at the conceptual level, but it will also be important to be clear on the ontological and epistemic assumptions that ground or support those conceptual understandings of what distinguishes that which is normatively relevant to 'politics', 'public law', or 'public interest' from what is relevant to 'morality'.

5.4. Theoretical Options and Biting Bullets

From the discussion above, it has hopefully become clear that more metaethical investigations are needed for the effective theorising of criminalisation from a 'political turn' perspective. As shown in the works of different theorists, a large part of the point of the 'political turn' in the first place is to support a normative distinction in the context of the criminal law, between those principles that are linked to political life and those that are not (which can be referred to as 'morality' as a whole or as something like 'inter-personal' morality). The point in showing this has not been, again, to agree or disagree with the distinction. Rather, my aim is to show that if theorists want the distinction to make sense, they will be better off by addressing the metaethics involved in making such a distinction.

There are, I propose, two possible options a theorist can take if they want to answer the question of distinctness between 'political' and 'moral'. Option 1, as I will call it, is to try and carve out a *part* of morality conceptually and argue that said portion of morality is distinctly political, in contrast to the rest of morality that will count as something else (such as 'inter-personal'). In other words, the distinctness of what we can call political morality, of its norms and

reasons, is not one of domain, but rather of something *about* the moral norms and principles. This is a subject-matter approach, to borrow the language of Chiao for different purposes than what he uses it for, and can be based on either the *target* of political morality or on the *source of authoritativeness* of political morality. If it is based on the target of political morality, then something will need to be said about what is special about this target: is it the amount of people to whom the moral facts apply? Is it the kinds of agents involved in the actions that are governed by this particular morality? And once this is determined, then it will be beneficial to discuss the metaethical nature of those targets: what properties of their existence give those targets their distinctiveness and how one goes about gaining knowledge about them, for example. The problem with this approach, however, is that as was briefly mentioned in chapter 2, trying to delimit the conceptual content of 'morality', let alone 'political morality', based on its target is not going to be a fruitful exercise. This is because the potential properties of the actions that would be the target of this morality can also plausibly be the target of other normative domains, and in this case also of the part of morality that is not political. Therefore, perhaps the more fruitful avenue will be to base the distinction within morality as one of the sources of authoritativeness, i.e. on the normativity of political morality. Again, this view will require us to make a lot of metaethical assumptions clear from the start – what authoritative normativity looks like, what ontological properties of the norms and reasons of political morality would provide that authoritativeness, and how we as decision-makers access those properties – but this avenue seems, at least at first glance, like a better path for an Option 1 theorist. This is because this kind of Option 1 theorist can make the argument that something about the political process – the fact that it is democratic, that it is liberal, that it is a republic, the fact that citizens consent to the process, or any other substantive political theory that the theorist might want to use – provides the normative

facts of political morality with a distinctive kind of authority; distinct enough to differentiate it from the rest of moral facts.[475] I will not delve deeper into this idea, but merely wish to signpost the possibility of this kind of argument for those theorists inclined to Option 1.

In contrast, Option 2 is a stronger claim and takes a different approach to Option 1. To follow Option 2, one would need to argue that there is a distinct normative *domain* for politics and political normativity; that this is *not* the same domain as morality. One way to do this would be to follow something along the lines of what has been dubbed 'political realism', whose proponents "maintain that political theory should begin (in a justificatory rather than temporal sense) not with the explication of moral ideals (of justice, freedom, rights, etc.) which are then taken to settle the questions of value and principle in the political realm but in a (typically interpretative) understanding of the practice of politics itself."[476] This view is particularly conscious of the actual practice of politics, and recognises that the way in which political decisions are made is importantly different to those made in the context of morality. As Raymond Geuss puts it, "politics as we know it is a matter of differential choice: opting for A *rather than* B. Thus politics is not about doing what is good or rational or beneficial *simpliciter*—it is not even obvious that that is an internally coherent thought at all—but about the pursuit of what is good in a particular concrete case by agents with limited powers and resources, where choice of one thing to pursue means failure to choose and pursue another."[477] Bernard Williams has also been dubbed a political realist, and his view of politics also looks closely at how political decisions are actually

[475] This might be one way of reading Raz's Pre-emptive thesis regarding the broader moral legitimacy of the authority of the state, discussed in Part I in Raz (n 224).
[476] Enzo Rossi and Matt Sleat, 'Realism in Normative Political Theory: What Is Political Realism?' [2014] Philosophy Compass 1, 2.
[477] Raymond Geuss, *Philosophy and Real Politics* (Princeton University Press 2008) 30–31.

made, and importantly at what the result of a political deliberation process conveys:

> "A very important reason for thinking in terms of the political is that a political decision – the conclusion of a political deliberation which brings all sorts of considerations, considerations of principle along with others, to one focus of decision – is that such a decision does not in itself announce that the other party was morally wrong or, indeed, wrong at all. What it immediately announces is that *they have lost*."[478]

It is important to keep in mind, however, that the potential distinctness of politics as a normative domain does not entail that other normative domains become irrelevant, or that one cannot still make normative statements *about* a domain from the perspective of another domain. So, for example, it is not irrational to say that a political action might be legitimate in a political context, yet still morally reprehensible. In other words, there is a difference between an action being politically legitimate or justified and the moral judgment that one may make *about* that political action – morality does not have a normative monopoly on legitimacy or justification. The point an Option 2 theorist could make about this issue is that, as Rossi and Sleat put it, "Morality may be successful at causally influencing politics as any other grand aspiration, but that is not to say that it can purport to guide political behaviour in the same way in which it guides personal conduct nor that it always provides a suitable blueprint for action in a sphere that is constituted by disagreement, contestation, and coercion."[479]

---

[478] Bernard Williams, *In the Beginning Was the Deed - Realism and Moralism in Political Argument* (Princeton University Press 2005) 13.
[479] Rossi and Sleat (n 476) 4.

Option 2, however, is obviously not an uncontroversial position to take. As Leader Maynard and Worsnip put it, "most moralist political theorists […] endorse some principles that are specific to politics, but hold that these principles still qualify as moral principles."[480] In turn, the metaethical commitments that might inform an Option 2 theorist and their understanding of political normativity, such as some form of constructivism which Leader Maynard and Worsnip think might be the case for some political realists,[481] will require a further metaethical explanation for why this constructivism makes sense for political normativity (and, additionally, for moral normativity if the theorist is also a constructivist for moral facts). Once again, my point in stating this difficulty is not to try to argue for or against Option 2, but instead to highlight the importance of addressing the metaethical implications of holding this kind of distinction.

Therefore, to summarise, if the distinction between the 'political' and the 'moral', in either its Option 1 or Option 2 form, is to be made intelligible, it would be beneficial to present the distinction with the metaethical implications of doing so in mind and discussing said implications explicitly in one's theorising. If one is an Option 1 theorist – as Chiao, Thorburn, and Brudner seem to be – then one's theory for evaluating the criminal law and criminalisation as distinctly 'political' entities would greatly benefit from discussing how, exactly, one carves out the conceptual arena for political morality. And if my suggestion were correct, and a plausible way of doing this might be to argue for a distinctive authoritativeness for political morality, then it behoves Option 1 theorists to explain what is the metaethical nature of that authoritativeness, and how a decision-maker can go about finding the politically-relevant facts that provide that authoritativeness. In turn, if one is

[480] Leader Maynard and Worsnip (n 435) 761.
[481] ibid 773–776.

an Option 2 theorist, then it is even more important to be explicit about the metaethical commitments involved in insisting that political normativity is distinct from morality as a normative domain.

In conclusion, when it comes specifically to criminalisation theorising, if a theorist wishes to put forward the distinction between the 'political' and the 'moral' as the basis for a criminalisation theory, they will have to choose which bullet they are willing to bite. On the one hand, Option 1 theorists will have to accept that their criminalisation theories are, in a sense, a different variety of legal moralist theory, rather than a more radical alternative to it, albeit being a distinct strand that focuses on political morality, and therefore one becomes a 'political moralist'. In this sense, an Option 1 theorist gains the possibility of using the language of wrongfulness to discuss both the declarations of the criminal law through criminalisation and the deliberation process for deciding what conducts to criminalise, but they will be able to directly point to political morality and argue that wrongfulness *in that specific sense* is the 'moral wrongfulness' relevant to criminalisation. Option 2 theorists, in contrast, will have to bite the bullet of having to be a 'political realist' and provide and engage with answers to conceptual, ontological, and epistemic questions about the 'political', in a way that theorists of the 'political turn' have not yet done. Either way, the 'political turn' in criminalisation theorising has its work cut out for itself.

CHAPTER 6 – CONCLUSION AND WHERE DO WE GO FROM HERE?

Let us now take stock of the ideas discussed in this thesis. I began by saying that our criminalisation theorising would be better off by informing it with metaethical investigations, particularly if we are to continue to develop normative theories of criminalisation that use wrongfulness as a theoretical tool. I argued that doing so avoids theorists talking past each other, as well as allowing us to fill theoretical gaps in our understanding of criminalisation, both as an act and as a decision-making process. In chapter 2, I provided some basic understandings of the main metaethical questions that are relevant to the use of moral wrongfulness as a normative tool, and posed the conceptual, ontological and epistemic questions, along with discussing some possible answers that have been given to these questions. Chapter 3 was one of my original contributions by presenting the speech-act model for understanding criminalisation, and I argued that any kind of wrongness that might be stated through the act of criminalising a conduct is best understood as an implicit assertion of said wrongfulness. I then applied the metaethical questions to this understanding of criminalisation, and showed that the way in which we propose the inferential process to be for understanding that implicit assertion of moral wrongness requires a metaethical explanation. Chapter 4 was a further original contribution by summarising some recent normative criminalisation theories, reconstructing what each theorist's deliberation process for criminalisation should look like, and constructed for each theorist answers to the metaethical questions posed. I showed that each theorist had slightly different metaethical commitments in their theorising, and that for potential decision-makers trying to apply the proposed deliberation processes, metaethical differences matter for determining the epistemic methodology that they will employ for their deliberation process. Finally, in chapter 5 I explored some of the more recent 'political turn'

criminalisation theories and showed that there, too, metaethical commitments are playing an important part in these kinds of theories. In particular, I argued that the distinction these theorists make between 'morality' and 'politics' is one with metaethical implications, but that none of these theorists are addressing them. I ended by suggesting ways forward in which these kinds of theorists can inform their views from metaethics and make sense of the distinction that they are interested in making.

As readers might have gleaned by now, this thesis has mostly been presented as a diagnostic for the state of the field in criminalisation theory. The tone I have gone for is mostly observational rather than imperative, and I have made my best attempts at staying as ecumenical as possible. Nonetheless, this thesis has shown that there are important theoretical gaps in our normative understanding and theorising of criminalisation, and it has provided examples of where those gaps lead to problems in criminalisation theories. In doing so, I am looking to provide a way forward for future criminalisation theorists, in order to show the importance of being aware and explicit about one's metaethical commitments, as well as putting forward the suggestion that our theorising will be better off by complementing it with metaethical investigations. Doing so, I argue, will allow us to provide better, metaethically informed normative criminalisation theories, that better explain what is being said when we criminalise, as well as how we should decide to say so. Ultimately, this should allow us to start filling the theoretical gaps that this thesis has endeavoured to point out.

If a criminalisation theory, then, is going to use moral wrongfulness as its main contender for explaining criminalisation – either as part of what is being said when a conduct is criminalised or as how we should decide what to

criminalised – it needs to be able to provide an answer to, among others, the following questions:

1. What does it mean for something to be morally wrongful and how is it different from other kinds of wrongfulness?
2. What are we looking for when referring to moral wrongfulness and how do we identify it?
3. How do we know when something is morally wrong and what do we need to do to know it?

Similarly, if a theorist wishes to do normative criminalisation theory that is not focused on moral wrongdoing but something else, such as the political turn theorists, I have aimed to show that they too would greatly benefit from engaging in metaethical investigations for their theorising. In essence, making sure that the normative theorising that one proposes fits with an account of how normativity fits with our understanding of our concepts, of our reality and of our knowledge is a useful endeavour for anyone who is interested in understanding criminalisation from a normative perspective.

My aim in this thesis has been to provide some theoretical tools that allow for the discussions taking place around criminalisation and its theorising to be somewhat more structured and organised in their analysis, and to begin to look for answers to the questions above. I have aimed at identifying what questions go where in the conversation, and to establish some concepts, models and issues that should make the conversation run, at least somewhat, more smoothly. In particular, I have tried to provide theoretical tools to better understand wrongfulness as a normative concept, and to bring attention onto the myriad of different possible meanings and contexts that can provide the content for our understandings of what 'wrong' is in the context of

criminalisation, and how best to go about using it as a concept that can be useful to our understanding of criminalisation.

There are two further questions which I wish to make explicit here, that I believe flow from the framework I have set up in this thesis regarding the use of moral wrongfulness as a tool for understanding criminalisation and which I must leave unanswered in this thesis. First, if moral wrongfulness is considered problematic from a metaethical point of view, then should we continue to use it in our discussions of criminalisation? Second, if we answer 'no' to the previous question, then what alternatives do we have?

The main take-away that I wish to leave readers with is the following. A criminalisation theory should be able to answer not only what is being said through a criminalisation claim and what principles should be used to deliberate on whether we should make a criminalisation claim. It should also give answers to the conceptual, ontological and epistemic questions posed in this thesis, if it is to use wrongfulness as a normative tool for the theory, and even more so if it wants to use moral wrongfulness specifically. In other words, a criminalisation theorist should be prepared to present a story about the metaethics and meta-normativity of their normative domain of choice, and explain how the conceptual, ontological and epistemic nature of the norms and reasons they appeal to affect how their theoretical models work.

So, in a way, my conclusion is just the first half of the steps needed to do what I propose here. It is intended as an indication of where to go from here in the context of theorising criminalisation. I have identified that we need the story: now we just need to tell it.

BIBLIOGRAPHY

- Allott N and Shaer B, 'The Illocutionary Force of Laws' (2018) 61 Inquiry 351
- Andenaes J, 'The Moral or Educative Influence of Criminal Law' (1971) 27 Journal of Social Issues 17
- Archer A, 'The Problem with Moralism' (2018) 31 Ratio 342
- Audi R, 'Moderate Intuitionism and the Epistemology of Moral Judgment' (1998) 1 Ethical Theory and Moral Practice 15
- Austin JL, *How To Do Things With Words* (Oxford University Press 1975)
- Ayer AJ, *Language, Truth and Logic* (Dover Publications 2012) <http://ebookcentral.proquest.com/lib/ed/detail.action?docID=1889619> accessed 29 June 2023
- Bach K and Harnish R, *Linguistic Communication and Speech Acts* (MIT 1979)
- Bagnoli C, *Constructivism in Ethics* (Cambridge University Press 2013) <http://ebookcentral.proquest.com/lib/ed/detail.action?docID=1303641> accessed 29 June 2023
- Baker D, 'Skepticism About Ought Simpliciter' (2018) 13 Oxford Studies in Metaethics
- ––, 'The Varieties of Normativity' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2018)
- Barker S, Crerar C and Goetze TS, 'Harms and Wrongs in Epistemic Practice' (2018) 84 Royal Institute of Philosophy Supplement 1
- Barry M, 'Constructivism', *The Routledge Handbook of Metaethics* (Routledge 2018)
- Berleant A, 'Negative Aesthetics and Everyday Life' (2011) 1
- Blackburn S, 'Circles, Finks, Smells, and Biconditionals □', *Practical Tortoise Raising* (Oxford University Press 2010)
- Boyd R, 'How To Be a Moral Realist' in Stephen Darwall, Allan Gibbard and Peter Railton (eds), *Moral Discourse & Practice - Some Philosophical Approaches* (Oxford University Press 1997)
- Brink DO, *Moral Realism and the Foundations of Ethics* (Cambridge

University Press 1989)

- Brown JM, 'Collateral Legal Consequences of Criminal Convictions in a Society of Equals' [2020] Criminal Law and Philosophy <https://doi.org/10.1007/s11572-020-09544-7>
- Brudner A, *Punishment and Freedom - A Liberal Theory of Penal Justice* (Oxford University Press 2009)
- Chiao V, 'What Is the Criminal Law For?' (2016) 35 Law and Philosophy 137
- ––, *Criminal Law in the Age of the Administrative State* (Oxford University Press 2019)
- Chrisman M, *What Is This Thing Called Metaethics?* (Taylor & Francis Group 2016) <http://ebookcentral.proquest.com/lib/ed/detail.action?docID=4747553> accessed 19 November 2022
- Colebrook R, 'The Irrationality of Folk Metaethics' (2021) 34 Philosophical Psychology 684
- Copp D, *Morality, Normativity & Society* (Oxford University Press 1995)
- ––, 'Why Naturalism?' (2003) 6 Ethical Theory and Moral Practice 179
- Cornford A, 'Rethinking the Wrongness Constraint on Criminalisation' (2017) 36 Law and Philosophy 615
- Darwall S, 'Ethics and Morality' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2018)
- Devlin P, *The Enforcement of Morals* (Oxford University Press 1965)
- Duff RA, *Punishment, Communication and Community* (Oxford University Press 2001)
- ––, *Answering for Crime - Responsibility and Liability in the Criminal Law* (Hart Publishing 2009)
- ––, 'Responsibility, Citizenship, and Criminal Law' in RA Duff and Stuart P Green (eds), *Philosophical Foundations of Criminal Law* (Oxford University Press 2011)
- ––, 'Towards a Modest Legal Moralism' (2012) 8 Criminal Law and Philosophy 217
- ––, *The Realm of Criminal Law* (Oxford University Press 2018)

- –– (eds), *Criminalization - The Political Morality of the Criminal Law* (Oxford University Press 2014)
- Dworkin R, 'Objectivity and Truth: You'd Better Believe It' (1996) 25 Philosophy and Public Affairs 87
- Edwards J, 'Coming Clean About the Criminal Law' (2011) 5 Criminal Law and Philosophy 315
- ––, 'Harm Principles' (2015) 20 Legal Theory 253
- ––, 'An Instrumental Legal Moralism' (2018) 3 Oxford Studies in Philosophy of Law: Volume 3
- Eklund M, *Choosing Normative Concepts* (Oxford University Press 2017)
- Enoch D, 'The Disorder of Public Reason' (2013) 124 Ethics 141
- ––, 'Non-Naturalistic Realism in Metaethics' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook in Metaethics* (Routledge 2019)
- Farmer L, 'Criminal Wrongs in Historical Perspective' in Duff and others (eds), *The Boundaries of the Criminal Law* (Oxford University Press 2010)
- ––, 'Criminal Law as an Institution' in RA Duff and others (eds), *Criminalization - The Political Morality of the Criminal Law* (Oxford University Press 2014)
- ––, *Making the Modern Criminal Law - Criminalization and Civil Order* (Oxford University Publlishing 2016)
- Feinberg J, *Doing and Deserving* (Princeton University Press 1970)
- ––, *Moral Limits of the Criminal Law* (Oxford University Press 1984)
- Finlay S, *Confusion of Tongues: A Theory of Normative Language* (Oxford University Press 2016)
- Fisher A, *Metaethics: An Introduction* (Taylor & Francis Group 2014) <http://ebookcentral.proquest.com/lib/ed/detail.action?docID=1886850> accessed 29 June 2023
- Fitzpatrick WJ, 'Ethical Non-Naturalism and Normative Properties' in Michael Brady (ed), *New Waves in Metaethics* (Palgrave Macmillan 2011)
- Fletcher G, *Dear Prudence - The Nature & Normativity of Prudential Discourse* (Oxford University Press 2021)

- Fletcher G and Ridge M (eds), *Having It Both Ways: Hybrid Theories and Modern Metaethics* (Oxford University Press 2014)
- Foot P, 'Morality as a System of Hypothetical Imperatives' (1972) 81 Philosophical Review 305
- Frankena WK, 'Main Trends in Recent Philosophy: Moral Philosophy at Mid-Century' (1951) 69 Philosophical Review 44
- Fumerton RA, *Reason and Morality* (Cornell University Press 1990)
- Gauthier D, *Morals by Agreement* (Oxford University Press 1987)
- Gendler Szabó Z and Thomason RH (eds), *Philosophy of Language* (Cambridge University Press 2019)
- Gerritsen E, 'Demistifying Normativity - Morality, Error Theory and the Authority of Norms' (PhD Thesis, Groningen/St Andrews/Stirling 2022) <https://research-repository.st-andrews.ac.uk/handle/10023/27003>
- Gert B and Gert J, 'The Definition of Morality', *Stanford Encyclopedia of Philosophy* (2020)
- Gert J, 'Normative Strength and the Balance of Reasons' (2007) 1 Philosophical Review
- Gettier EL, 'Is Justified True Belief Knowledge?' (1963) 23 Analysis 121
- Geuss R, *Philosophy and Real Politics* (Princeton University Press 2008)
- Gibbard A, *Wise Choices, Apt Feelings* (Oxford University Press 1990)
- Gilbert J, 'Moral Inescapability' (1975) 10 Journal of Thought 8
- Goldsworthy J, 'Externalism, Internalism and Moral Scepticism' (1992) 70 Australasian Journal of Philosophy 40
- Grice HP, *Studies in the Way of Words* (Harvard University Press 1991) <http://hdl.handle.net/2027/heb08428.0001.001> accessed 21 October 2022
- Grice P, 'Further Notes on Logic and Conversation' in Peter Cole (ed), *Pragmatics* (Academic Press 1978)
- Gundersen EB, 'Making Sense of Response-Dependence' (University of St Andrews 2006)
- Habermas J, *Moral Consciousness and Communicative Action* (Polity Press 1990)

- Hart HLA, *Law, Liberty and Morality* (Oxford University Press 1963)
- ––, *Punishment and Responsibility - Essays in the Philosophy of Law* (Oxford University Press 2008)
- ––, *The Concept of Law* (3rd edn, Oxford University Press 2012)
- Hoskins Z, 'Criminalization and the Collateral Consequences of Conviction' (2018) 12 Criminal Law and Philosophy 625
- Husak D, *Overcriminalization: The Limits of the Law* (Oxford University Publlishing 2008)
- ––, *The Philosophy of Criminal Law* (Oxford University Press 2010)
- ––, 'Wrongs, Crimes, and Criminalization' (2019) 13 Criminal Law and Philosophy 393
- Husak D and de Marneffe P, *The Legalization of Drugs* (Cambridge University Press 2005)
- Hussain NJZ and Shah N, 'Misunderstanding Metaethics: Korsgaard's Rejection of Realism' in Russ Shaffer-Landau (ed), *Oxford Studies in Metaethics*, vol 1 (Oxford University Press 2006)
- Johnston M, 'Are Manifest Qualities Response-Dependent?': (1998) 81 Monist 3
- Joyce R, *The Evolution of Morality* (MIT 2006)
- ––, 'Moral Anti-Realism', *Stanford Encyclopedia of Philosophy* (2016)
- Kagan S, *Normative Ethics* (Westview Press 1998)
- Kamm F, *Intricate Ethics: Rights, Responsibilities, and Permissible Harm* (Oxford University Press 2007)
- Kiesewetter B, *The Normativity of Rationality* (Oxford University Press 2017)
- Kirkham R, *Theories of Truth, a Critical Introduction* (MIT 1992)
- Kleinfeld J, 'Embodied Ethical Life and the Criminal Law' in Chad Flanders and Zachary Hoskins (eds), *The New Philosophy of Criminal Law* (Rowman & Littlefield International 2016)
- Koller P, 'On the Nature of Norms: On the Nature of Norms' (2014) 27 Ratio Juris 155
- Korsgaard C, *The Sources of Normativity* (Cambridge University Press 1996)
- Korsgaard CM, *The Constitution of Agency: Essays on Practical Reason and*

Moral Psychology (Oxford University Press 2008) <https://doi.org/10.1093/acprof:oso/9780199552733.003.0008> accessed 29 June 2023

- Lacey N, 'Historicising Criminalisation: Conceptual and Empirical Issues' (2009) 72 Modern Law Review 936
- Lacey N and Pickard H, 'Why Standing to Blame May Be Lost but Authority to Hold Accountable Retained: Criminal Law as a Regulative Public Institution' (2021) 104 The Monist 265
- Leader Maynard J and Worsnip A, 'Is There a Distinctively Political Normativity?' (2018) 128 Ethics 756
- Leiter B (ed), *Objectivity in Law and Morals* (Cambridge University Press 2001)
- Lutz M and Lenman J, 'Moral Naturalism', *The Stanford Encyclopedia of Philosophy* (Spring 2021, Metaphysics Research Lab, Stanford University 2021) <https://plato.stanford.edu/archives/spr2021/entries/naturalism-moral/> accessed 28 June 2023
- Lutz M and Ross J, 'Moral Skepticism' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2018)
- Mackie JL, *Ethics - Inventing Right and Wrong* (Penguin Publishing 1977)
- Marmor A, *The Language of Law* (Oxford University Press 2014)
- Marshall SE and Duff RA, 'Criminalization and Sharing Wrongs' in Paul H Robinson, Stephen P Garvey and Kimberly Kessler Ferzan (eds), *Criminal Law Conversations* (2009)
- McHugh C, Way J and Whiting D (eds), *Normativity: Epistemic and Pragmatic* (Oxford University Press 2018)
- McPherson T, 'Metaethics and the Autonomy of Morality' (2008) 8
- ––, 'Against Quietist Normative Realism' (2011) 154 Philosophical Studies 223
- ––, 'Unifying Moral Methodology' (2012) 93 Pacific Philosophical Quarterly 523
- McPherson T and Plunkett D, 'The Nature and Explanatory Ambitions of Metaethics' in Tristram McPherson and David Plunkett (eds), *The Routledge*

Hanbook of Metaethics (Routledge 2018)

- Miller A, *Contemporary Metaethics - An Introduction* (2nd edn, Polity Press 2013)
- Moore MS, 'Moral Reality Revisited' (1992) 90 Michigan Law Review 2424
- ––, *Placing Blame - A General Theory of the Criminal Law* (Oxford University Press 1997)
- ––, 'A Tale of Two Theories' (2009) 28 Criminal Justice Ethics 27
- Morris CW, 'A Contractarian Account of Moral Justification' in Walter Sinnot-Armstrong and Mark Timmons (eds), *Moral Knowledge? New Readings in Moral Epistemology* (Oxford University Press 1996)
- Mullock P, 'The Logic of Entrapment' (1985) 46 University of Pittsburgh Law Review 739
- Myers RH, 'The Inescapability of Moral Reasons' (1999) 59 Philosophy and Phenomenological Research 281
- Olson J, *Moral Error Theory - History, Critique, Defence* (Oxford University Press 2014)
- ––, 'Error Theory in Metaethics' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2019)
- Parfit D, *Reasons and Persons* (1st edn, Oxford University PressOxford 1986) <https://academic.oup.com/book/12484/chapter/163161485> accessed 26 June 2023
- Pereboom D, *Free Will, Agency, and Meaning in Life* (Oxford University Press 2014)
- Petersen TS, *Why Criminalize? New Perspectives on Normative Principles of Criminalization*, vol 134 (Springer International Publishing 2020) <http://link.springer.com/10.1007/978-3-030-34690-4> accessed 16 May 2023
- Plunkett D, 'Robust Normativity, Morality and Legal Positivism' in David Plunkett, Scott J Shapiro and Kevin Toh (eds), *Dimensions of Normativity - New Essays on Metaethics and Jurisprudence* (Oxford University Press 2019)
- Pölzler T and Wright JC, 'Anti-Realist Pluralism: A New Approach to Folk Metaethics' (2020) 11 Review of Philosophy and Psychology 53

- Portmore DW, 'Are Moral Reasons Morally Overriding?' (2008) 11 Ethical Theory and Moral Practice
- Railton P, 'Moral Realism: Prospects and Problems' in Walter Sinnot-Armstrong and Mark Timmons (eds), *Moral Knowledge? New Readings in Moral Epistemology* (Oxford University Press 1996)
- Rawls J, *A Theory of Justice* (Harvard University Press 1971)
- Raz J, *The Morality of Freedom* (Oxford University Press 1986)
- ––, *The Authority of Law, Essays on Law and Morality* (2nd edn, Oxford University Press 2009)
- Renzo M, 'Responsibility and Answerability in the Criminal Law' in RA Duff and others (eds), *The Constitution of the Criminal Law* (2013)
- Ridge M, 'Reasons for Action: Agent-Neutral vs. Agent-Relative', *Stanford Encyclopedia of Philosophy* (2017)
- ––, 'Moral Non-Naturalism', *The Stanford Encyclopedia of Philosophy* (Fall 2019, Metaphysics Research Lab, Stanford University 2019) <https://plato.stanford.edu/archives/fall2019/entries/moral-non-naturalism/> accessed 28 June 2023
- Rosati CS, 'Mind-Dependence and Moral Realism' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2017)
- Rossi E and Sleat M, 'Realism in Normative Political Theory: What Is Political Realism?' [2014] Philosophy Compass 1
- Sayre-McCord G, 'Coherentist Epistemology and Moral Theory' in Walter Sinnot-Armstrong (ed), *Moral Knowledge? New Readings in Moral Epistemology* (Oxford University Press 1996)
- Scanlon TM, 'Contractualism and Utilitarianism', *Utilitarianism and Beyond* (Cambridge University Press 1982)
- Scanlon TM, *What We Owe to Each Other* (Harvard University Press 1998) <http://www.jstor.org/stable/j.ctv134vmrn> accessed 29 June 2023
- Schonsheck J, *On Criminalization* (Kluwer Law International 1994)
- Schroeder M, 'Hybrid Expressivism: Virtues and Vices' (2009) 119 Ethics 257
- Searle J, *Speech Acts - An Essay in the Philosophy of Language* (Cambridge

University Press 1969)

- ––, *Expression and Meaning - Studies in the Theory of Speech Acts* (Cambridge University Press 1979)
- Simester AP and Von Hirsch A, *Crimes, Harms and Wrongs - On the Principles of Criminalisation* (Hart Publishing 2011)
- Sinnott-Armstrong W, 'Moral Skepticisms and Justification' in Walter Sinnott-Armstrong and Mark Timmons (eds), *Moral Knowledge? New Readings in Moral Epistemology* (Oxford University Press 1996)
- ––, *Moral Skepticisms* (Oxford University Press 2006)
- ––, 'The Disunity of Morality' in S Matthew Liao (ed), *Moral Brains: The Neuroscience of Morality* (Oxford University Press 2016)
- Smith M, *The Moral Problem* (Blackwell Publishers 1994)
- Srinivasan A, 'Feminism and Metaethics' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2018)
- Street S, 'Constructivism about Reasons' in Russ Shafer-Landau (ed), *Oxford Studies in Metaethics*, vol 3 (Oxford University Press 2008)
- ––, 'What Is Constructivism in Ethics and Metaethics?' (2010) 5 Philosophy Compass 363
- Tadros V, 'The Architecture of Criminalization' (2009) 28 Criminal Justice Ethics 74
- ––, *The Ends of Harm - The Moral Foundations of Criminal Law* (Oxford University Press 2011)
- ––, 'Wrongness and Criminalization' in Andrei Marmor (ed), *The Routledge Companion to Philosophy of Law* (Routledge 2012)
- ––, *Wrongs and Crimes* (Oxford University Press 2016)
- ––, 'Responses to Wrongs and Crimes' (2019) 13 Criminal Law and Philosophy 455
- Thorburn M, 'Criminal Law as Public Law' in RA Duff and Stuart P Green (eds), *Philosophical Foundations of Criminal Law* (Oxford University Press 2011)
- ––, 'Constitutionalism and the Limits of the Criminal Law' in RA Duff and others (eds), *The Structures of the Criminal Law* (Oxford University Press

2012)

- Tiffany E, 'Deflationary Normative Pluralism' (2007) 33 Canadian Journal of Philosophy Supplementary Volume 231
- Tomlin P, 'Duffing Up the Criminal Law?' (2019) 14 Criminal Law and Philosophy 319
- Tropman E, 'Intuitionism in Moral Epistemology' in Tristram McPherson and David Plunkett (eds), *The Routledge Handbook of Metaethics* (Routledge 2018)
- Van Hoecke M, *Law as Communication* (Hart Publishing 2002)
- Višak T, 'Sacrifices of Self Are Prudential Harms: A Reply to Carbonell' (2015) 19 The Journal of Ethics 219
- Wilenmann J, 'Framing Meaning through Criminalization: A Test for the Theory of Criminalization' (2019) 22 New Criminal Law Review
- Williams B, *In the Beginning Was the Deed - Realism and Moralism in Political Argument* (Princeton University Press 2005)
- ––, *Ethics and the Limits of Philosophy* (Routledge 2011)
- Wong DB, 'Relativism and Pluralism in Moral Epistemology' in Aaron Zimmerman, Karen Jones and Mark Timmons (eds), *The Routledge Handbook of Moral Epistemology*