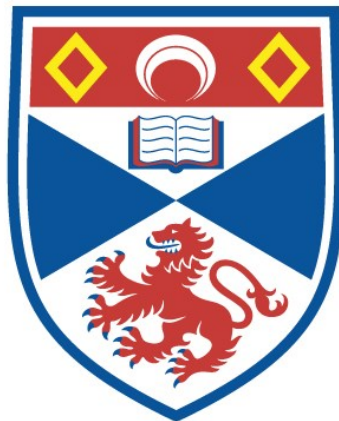


Hume's theory of obligation: morality between nature and artifice

Xiao Qi

A thesis submitted for the degree of PhD
at the
University of St Andrews



2025

Full metadata for this item is available in
St Andrews Research Repository
at:

<https://research-repository.st-andrews.ac.uk/>

Identifier to use to cite or link to this thesis:

DOI: <https://doi.org/10.17630/sta/1227>

This item is protected by original copyright

Candidate's declaration

I, Xiao Qi, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 80,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in September 2020.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 19 December 2024

Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date 18 December 2024

Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Xiao Qi, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Electronic copy

No embargo on electronic copy.

Date 19 December 2024

Signature of candidate

Date 19 December 2024

Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Xiao Qi, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date 19 December 2024

Signature of candidate

Abstract

Hume's account of obligation has been widely discussed in the literature, but many aspects of his conception of obligation, including its definition, taxonomy, and scope, are still subjects of interpretative controversy. Some scholars read his concept of moral obligation as denoting a determining motive in regard to moral values, while some read it as identical to the approbation of virtue which cannot motivate by itself. I argue against both views in this thesis. The primary concern of Hume's account of obligation, in my reading, is not with motivation but with accommodating moral obligation in a virtue-centred moral theory. For him, the sentiment of moral obligation is a specific form of disapprobation, and it has different features when felt about 'natural virtues' like benevolence and 'artificial virtues' like justice. Noteworthily, for justice, considerations about law and obligation precede those about character, which makes justice an anomalous case for Hume's virtue-centred theory, and he develops an account of the conventional origin of justice to explain its law-regulated features. The idea of obligation, on this account, arises from principles in human nature in a social and historical process but has features in tension with such principles' natural inclinations.

In Chapter 1, I delineate the basic features of Hume's conception of obligation from both positive and negative perspectives. Chapter 2 discusses his taxonomy of obligation in the light of its relationship to natural law theories. Chapter 3 examines the nature, origin, and scope of the sentiment of moral obligation. Chapter 4 explores how concerns about features of the obligation to justice motivate Hume to propose a distinction between natural and artificial virtues. Then, Chapter 5 discusses how Hume constructs a genealogical account to explain the anomalous features of justice, and Chapter 6 explores whether and how this account can do its job.

Acknowledgement

Writing this thesis has been a fascinating adventure for me. Even by the time I only had several paragraphs in the conclusion left to be finished, I could not believe this is what I could achieve. All of this would be impossible without the kind support I have been receiving from my guides and companions on this journey. Out of all the doubts, struggles, and hesitations, their help has been my source of courage, confidence, and power to carry on.

First of all, I would like to give my heartfelt gratitude to James Harris, my primary supervisor. James's works were among those which, at the very beginning of my acquaintance with Hume, attracted me to Hume's theory of justice and obligation, from where this adventure started. For all these years, James has been an incredibly supportive supervisor, a meticulous reader and reviewer, a source of important and enlightening questions, and also a model for good research on the history of philosophy and good work on the position of a philosopher. My sincere gratitude also goes to Alexander Douglas, my secondary supervisor. I benefited a lot from our inspiring discussions, and Alex's advice has always been helpful and efficacious in locating weaknesses in my works and improving them. I would like to give special thanks to both my supervisors' patience and kindness as a student who has never been good at controlling the length of my works and catching up with deadlines. Their approval and encouragement have been the strongest source of my strength on this journey. I am very grateful to Aaron Garrett and Benjamin Sachs-Cobbe for offering to work as the examiners of my thesis; I owe them a lot for taking the time to read this work in a time around the holidays, and for making it possible for me to have my viva at a date perfectly suited to my schedule.

It is a great opportunity for me to be able to pursue my doctoral study here at St Andrews, which provides an exemplar for a lively, inclusive, and supportive academic environment for bolstering curiosity and sharing ideas. I feel very lucky to have benefitted from conversations here with Knud Haakonssen, Magaret Hampson, Mara van der Lugt, and my annual reviewers: Adam Etinson, Joseph Millum, and Jens Timmermann. I am also grateful to all members of the Philosophy Department for the amazing events from which I learned, and to Frédérique Stintzy for her help with my reading French courses. I owe my special thanks to my fantastic colleagues and friends, Enrico Galvagni and David Harmon, for all the thought-provoking reading groups and lovely pub hours, for their always being ready for help whenever I am in need, and for their

providing me with role models as fellow Ph.D. students. I will always miss the time when we philosophized together.

The development of my thoughts during my doctoral research has also benefited a lot from other scholars. My commentators and mentors at the Hume Society Conferences in Provo and Oxford: Donald Ainslie, Åsa Carlson, Rachel Cohon, and Elizabeth Radcliffe, provided me with invaluable comments and advice. I also learned a lot from discussions with Ruth Boeker, James Chamberlain, Michael Gill, Ryu Susato, and many other brilliant scholars at these conferences. I am very grateful for the ‘First Friday’ workshops in early modern philosophy at the University of Edinburgh, for its organizers, Jonathan Cottrell and Jennifer Smalligan Marušić, and its attendants, for providing me with an amazing occasion for sharing my research and getting acquainted with a group of excellent researchers. I should not omit friends who have given me precious feedbacks and helped me keep my philosophical ideas vivid and strong through online workshops and reading groups, especially through the Covid years: Canpu Chen, Jiahui Feng, Huimin Liu, Qianqian Wang, Erfan Xia, Qinglu Xie, Yi Yang, Yutao Zhao, Shuai Zhang, Hongbin Zhao, and all participants of three reading groups organized by Kaicheng Fang, Jiaxian Liu, and myself. I would like to thank them for giving me a special sense of community across far-away physical distances.

Let me also thank my fantastic friends in St Andrews: Kuoshi Hu, my roommate and the best Chinese and Korean chef in town; Zhuangwei Li, Hongjian Tan, and Tiancheng Yu, my old friends whom I am lucky enough to do my Ph.D. together with; and Jingyi Chen, Xuejia Ke, Jiangmei Liu, Yuhang Wang, and Mi Zhou, who have been sources of the brightest laughter even in the gloomiest days. Let me not forget the East Sand in St Andrews, whose beauty and tranquillity never fail to absorb all my negative emotions and make me feel alive again.

Finally, let me give my sincerest gratitude to my parents and parents-in-law, who gave me the firmest support in the world. My biggest thank goes to Fanhao Meng, my boyfriend when I started my Ph.D. who later became my husband, who, for almost fourteen years, has always been the most constant cause of my joy, love, and pride.

Funding

This work was supported by the China Scholarship Council-St Andrews Joint Scholarship.

Abbreviations

Works of David Hume:

- [T] *A Treatise of Human Nature*. Edited by L. A. Selby-Bigge. Revised by P. H. Nidditch. 2nd ed. Oxford: Clarendon Press, 1978.
- [EHU] *An Enquiry Concerning Human Understanding*, in *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. Edited by L. A. Selby-Bigge. Revised by P. H. Nidditch. 3rd ed. Oxford: Clarendon Press, 1975.
- [EPM] *An Enquiry Concerning the Principles of Morals*, in *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. Edited by L. A. Selby-Bigge. Revised by P. H. Nidditch. 3rd ed. Oxford: Clarendon Press, 1975.
- [DP] *A Dissertation on the Passions*, in *A Dissertation on the Passions and The Natural History of Religion*. Edited by Tom L. Beauchamp. Oxford: Clarendon Press, 2007.
- [E] *Essays Moral, Political, and Literary*. Edited by Eugene F. Miller. Indianapolis: Liberty Fund, 1987.
- [L] *The Letters of David Hume*. Edited by J. Y. T. Greig. 2 vols. New York: Oxford University Press, 2011.

Works of other authors:

Francisco Suárez

- [*Tractatus*] *A Treatise on Law and God the Lawgiver*, in *Selections from Three Works*. Edited by Thomas Pink. Translated by Gwladys L. Williams, Ammi Brown, John Waldron, and Henry Davis. Indianapolis: Liberty Fund, 2015.

Hugo Grotius

- [DJBP] *The Rights of War and Peace*. Edited by Richard Tuck. Translated by John Morrice. 3 vols. Indianapolis: Liberty Fund, 2005.

Thomas Hobbes

- [*De Cive*] *On the Citizen*. Edited by Richard Tuck and Michael Silverthorne. New York: Cambridge University Press, 1997.
- [*Leviathan*] *Leviathan*. Edited by Edwin Curley. Indianapolis and Cambridge: Hackett Publishing Company, 1994.

Richard Cumberland

- [DLN] *A Treatise of the Laws of Nature*. Edited by Jon Parkin. Translated by John Maxwell. Indianapolis: Liberty Fund, 2005.

Samuel Pufendorf

- [DJNG (1729)] *Of the Law of Nature and Nations*. Translated by Basil Kennett. 4th ed. London: J. Walthoe, R. Wilkin, J. and J. Bonwicke, S. Birt, T. Ward, and T. Osborne, 1729.
- [DJNG] *Of the Law of Nature and Nations*, in *The Political Writings of Samuel Pufendorf*. Edited by Craig L. Carr. Translated by Michael L. Seidler. New York; Oxford: Oxford University Press, 1994.
- [*De Officio*] *On the Duty of Man and Citizen According to Natural Law*. Edited by James Tully. Translated by Michael Silverthorne. Cambridge: Cambridge University Press, 1991.

Bernard Mandeville

- [*Fable*] *The Fable of the Bees, or Private Vice, Publick Benefits*. Edited by F. B. Kaye. 2 vols. Indianapolis: Liberty Fund, 1988.

Francis Hutcheson

- [*Inquiry*] *An Inquiry into the Original of Our Ideas of Beauty and Virtue*. Edited by Wolfgang Leidhold. 2nd ed. Indianapolis: Liberty Fund, 2008.
- [*Essay with Illustrations*] *An Essay on the Nature and Conduct of the Passions and Affections, with Illustrations on the Moral Sense*. Edited by Aaron Garrett. Indianapolis: Liberty Fund, 2002.
- [*Short Introduction*] *Philosophiae Moralis Institutio Compendiaria, with A Short Introduction to Moral Philosophy*. Edited by Luigi Turco. Indianapolis: Liberty Fund, 2007.

[*System*] *A System of Moral Philosophy*. 3 vols. Cambridge Library Collection
Philosophy. Cambridge: Cambridge University Press, 2014.

Henry Home, Lord Kames

[*Principles of Morality*] *Essays on the Principles of Morality and Natural Religion*. Edited by Marie
Catherine Moran. Indianapolis: Liberty Fund, 2005.

Adam Smith

[TMS] *The Theory of Moral Sentiments*. Edited by D. D. Raphael and A. L. Macfie.
Indianapolis: Liberty Fund, 1976.

Table of Contents

Abstract.....	1
Acknowledgement	6
Abbreviations.....	8
Chapter 1 Introduction: What Is (Not) Hume’s Theory of Obligation	13
1.1 Features of Hume’s Conception of Moral Obligation.....	14
1.2 Obligation and Motivation.....	18
1.3 Obligation and Virtue.....	25
1.4 Chapter Preview.....	33
Chapter 2 Natural Law and Hume’s Taxonomy of Obligation.....	37
2.1 Hume’s Taxonomy of Obligation.....	40
2.2 ‘Natural’ and ‘Moral’ Obligations.....	46
2.3 ‘Natural’ and ‘Artificial’ Obligations.....	55
2.4 Conclusion	66
Chapter 3 Unpacking the Sentiment of Moral Obligation.....	68
3.1 Obligation, Virtue, and Vice: A Preliminary Analysis	69
3.2 Nature and Expectation: The Origin of Moral Obligation.....	74
3.3 The Objects and Scope of Moral Obligation	81
3.4 Conclusion	91
Chapter 4 Moral Obligation and the Artificiality of Justice	92
4.1 Obligation, Virtue, and Law: <i>Treatise</i> 3.2.6.....	93

4.2 Obligation, Virtue, and Motive: <i>Treatise</i> 3.2.1	99
4.3 The Circle, the Sophistry, and the State of Nature.....	108
4.4 Conclusion	115
Chapter 5 Sentiment and Convention: Hume’s Method for Artificial Morality	117
5.1 Sentiment and History: Hume’s Two Questions about Justice	119
5.2 History and Reason: Hume’s Twofold Account of the Origin of Property	127
5.3 Reason and Custom: The ‘Real’ History of Justice	136
5.4 Conclusion	141
Chapter 6 From Interest to Morality: Convention and Obligation.....	143
6.1 Natural Obligation, Artificial Interest: The Conventional Origin of Obligation	145
6.2 Custom and Sentiment: From Natural to Moral Obligation	152
6.3 Law, Obligation, and Virtue: Replies to the Artificiality Arguments.....	163
6.4 Conclusion	172
Chapter 7 Conclusion: Morality between Nature and Artifice	174
Bibliography	180

Chapter 1

Introduction: What Is (Not) Hume's Theory of Obligation

David Hume's *account* of obligation have been widely discussed in the literature, although his 'theory of obligation' is an area which is, to a considerable extent, underexplored. His account of duty and obligation is widely discussed because it makes up a part of several of his most famous arguments in moral and political philosophy. These concepts are there in his arguments against moral rationalists, in his distinction between 'is' and 'ought', in his arguments for the thesis that justice is an 'artificial' instead of 'natural' virtue, in his account of the conventional origin of property and promises, and in his arguments against social contract theorists regarding the duty of political obedience. At the same time, Hume leaves too many things about duty and obligation unclarified. He does not provide any official *definitions* for these concepts. He does not provide a *catalogue* for duties. It appears that he has a *taxonomy* for different types of duties and obligations, but he does not explain on what ground his distinctions are made. Although Hume talks a lot about obligation, accordingly, it may fairly be questioned whether he has a 'theory of obligation' at all. Probably as a result, not many scholarly discussions have been dedicated to Hume's conception of obligation *as such*¹ despite the huge amount of controversy on the interpretation of his arguments in which the topic of obligation is involved.

The subject of this thesis is Hume's theory of obligation. If it cannot be assumed in advance that Hume has a 'theory of obligation', this thesis is an attempt to put Hume's account of obligation in a more coherent and systematic framework, locate it in the landscape of his moral philosophy, and in so doing, to work out what Hume's view on obligation is. This thesis is mainly interpretative and does not intend to evaluate to what extent Hume's view on obligation is compelling or attractive from the perspective of contemporary philosophy. The latter question is for another inquiry. Before that, it is worthwhile to work out what this view *is*. In this introductory chapter, I will start with a brief overview of three key features of Hume's conception of moral obligation. Then, I will discuss two perspectives from which Hume's account of obligation is usually examined by scholars, namely, obligation as a determining

¹ For noteworthy exceptions, see Wand (1956), Haakonssen (1978), Shaver (1992), Brown (1994), Darwall (1995, chap. 10), and Harris (2010a).

motive, and as an evaluative attitude. I will argue in favour of the second approach and suggest that Hume's primary concern with the topic of obligation is to give it a proper place in a virtue-centred theory.

1.1 Features of Hume's Conception of Moral Obligation

What, then, does Hume say about moral obligation at a broad level? When I use the term 'obligation' without qualification, I mean '*moral* obligation'. Hume talks about another type of obligation, i.e., 'natural obligation', which denotes a type of sense of interest; I will always refer to it as 'natural' or 'interested obligation' to avoid ambiguity. Hume's conception of moral obligation is, first, *sentimentalist* in the sense that judgements about obligation, like those about virtue and vice, are derived from moral sentiments. As Hume suggests in a passage in *A Treatise of Human Nature* which is the text closest to a 'definition' of moral obligation:

All morality depends upon our sentiments; and when any action, or quality of the mind, pleases us *after a certain manner*, we say it is virtuous; and when the neglect, or non-performance of it, displeases us *after a like manner*, we say that we lie under an obligation to perform it. (T 3.2.5.4; 517)

The *sentiment of moral obligation*, accordingly, is an unpleasant moral sentiment felt about the absence or omission of an action or mental quality (that is, in this context, a trait of character). We can derive from this type of sentiment, which is, in Hume's terms, an impression, the *idea* that one 'lies under an obligation' to perform the action or possess the mental quality at issue. Hume uses 'the sentiment of moral obligation' and 'the sense of duty' as synonyms² and does not posit any technical distinction between 'duty' and 'obligation'. For convenience, I will save the term 'duty' for the action or mental quality which one is morally obliged to perform or have and the term 'obligation' for the corresponding moral sentiment and the idea derived from this sentiment, though I do not presume that this is a distinction that Hume adopts.

The second key feature of Hume's view on duty and obligation is that these concepts are *non-fundamental* to his moral theory.³ Hume's main concern in his moral studies is with what

² See, for instance, T 3.2.2.8-9, 3.2.9.3; 479, 551.

³ It would be anachronistic to ascribe to Hume any version of the post-Kantian distinction between 'ethics' and 'morals', in the sense that morals concern what is required or binding on all rational persons (that is, in a Kantian sense, what is obligatory) while ethics denotes a domain either wider or narrower than the latter. Morality, as is examined by Hume in Book 3 of the *Treatise* and the second *Enquiry*, concerns all qualities of persons and actions to which moral sentiments respond: virtue and vice, right and wrong, as well as obligation; this is the meaning of 'moral' I adopt in this thesis. Hume also uses 'moral' in other senses like what concerns human nature or what concerns matters of facts (see, for instance, EHU 1.1, 4.18; 3, 35).

1.1 Features of Hume's Conception of Moral Obligation

is good and evil, virtuous and vicious, and he organizes his two major works on morality, Book 3 of the *Treatise* and the *Enquiry Concerning the Principles of Morals*, as a framework of different types of virtues and the principles according to which we morally approve of them. Duty and obligation are, as I will argue, derivative moral concepts in the sense that the sentiment of moral obligation is causally dependent on the approbation of virtue or disapprobation of vice, and for Hume, considerations of obligation make up a part of the moral domain but do not dominate it. Many scholars today read Hume's moral theory as a version of virtue ethics understood as a strand of moral theory opposed to deontology and consequentialism.⁴ For those who endorse this interpretation, as Rachel Cohon (2008b, 212–14) suggests, Hume's treatment of moral obligation provides a noteworthy example about how to give this concept a proper place in a virtue ethical theory. For those who doubt whether the virtue ethical interpretation captures Hume's view accurately, as James Harris (2010a, 25–33) does, Hume's account of obligation serves as a ground for disagreement. Annette Baier (2010, 124) argues that Hume's moral theory takes a middle ground between Aristotelian virtue theories and natural law and Kantian theories because, on one hand, it “secularizes and demystifies the concepts of obligation and of authority”, detaching them from divine command and a ‘mysterious’ (from an empirical perspective) idea of autonomous free will in the nominal world, and on the other hand, it “does this in a nonreductive distinction-preserving way”, keeping a place for moral obligation without confounding it with mere discouragement of vices or reducing it to interest and voluntary agreements. I agree with Baier's judgement in both aspects. I will argue that for Hume, moral obligation arises as a more complicated form of moral evaluation compared to those about virtue and vice as the former builds upon the latter with considerations about the general expectations about characters and codes of actions in a community, and I will come back to the relationship between virtue and obligation later in this chapter.

The most intriguing thesis in Hume's account of obligation, thirdly, is his claim that the idea of obligation is ‘*artificial*’ or dependent on social conventions (T 3.2.2.11; 491). Hume's account of obligation is closely intertwined with his theory of ‘artificial virtues’, especially that of justice understood as the disposition to observe the ‘three fundamental laws of nature’, that is, “*that of the stability of possession, of its transference by consent, and of the performance of*

⁴ For some representatives, see Hursthouse (1999), Greco (2013), Swanton (2015, chaps 3–5).

promises” (T 3.2.6.1; 526).⁵ In the *Treatise*, almost all of Hume's positive remarks on duty and obligation are contained in the part on artificial virtues⁶; the situation of *An Enquiry Concerning the Principles of Morals* is similar⁷, though the latter contains more discussions of duties which correspond to so-called ‘natural virtues’ like benevolence. Based on this close relation, Baier (2010, 130–35) suggests that Hume limits obligation to the domain of artificial virtues. More scholars are in favour of a more moderate reading, maintaining that Hume does acknowledge a moral obligation to have some natural virtues.⁸ He pays more attention to obligation in his theory of justice because this topic occupies a more significant place there. Jerome Schneewind (1998, 365–66) argues that Hume's distinction between natural and artificial virtues can be read as an attempt to explain the distinction between perfect and imperfect duties, passed down from Hugo Grotius and Samuel Pufendorf, in the context of a virtue-centred theory. It is questionable how accurate the correspondence between these distinctions is, although I agree with Schneewind that the need to accommodate perfect duties into his virtue-centred theory motivates Hume's distinction between natural and artificial virtues in the *Treatise*. He needs an account of human artifice to explain the origin of such rights and obligations since our judgements on these moral qualities are made in a manner in tension with the natural inclinations of our moral sentiments: these judgements are rule-directed, insensitive to particular circumstances, and precisely applied to objects and actions, while the approval of virtues and disapproval of vices are often not. Hume would probably agree with Bernard Williams (2006, 182–85) that obligation is only “one kind of ethical considerations among others” which make up a “particular institution” to secure people's mutual reliance in a sociable life, and Hume's key concern with his account of duty and obligation in the *Treatise* is to explain how such an institution emerged from human beings’

⁵ In Hume's moral works, the term ‘justice’ is used in at least three different senses: the narrowest merely pertains to the observation of the rules concerning the stability of external possessions or the rules of property; the second-narrow meaning pertains to all three ‘fundamental laws of nature’; and the widest meaning, which is not precisely delineated, may be closer to the common-life use of this term as covering rules that forbid injury against persons, rules of fairness, criminal laws, and the laws of nations. It appears that Hume did not bother to examine whether, and on what basis, these meanings of justice can be unified under one general definition or principle, and he felt free to switch between different meanings in different contexts. When talking about Hume's theory of justice in this thesis, I adopt the second-narrow meaning by default as I take it as most frequently employed by Hume in relevant contexts, and I will make notes if, to follow Hume's texts more closely, I use this term in other senses. For Hume's different uses of ‘justice’, especially see Baier (2010, chap. 4).

⁶ There is only one mention of obligation (T 3.3.1.25; 587) and two mentions of duty (T 3.3.1.27, 3.3.2.11; 590, 598) in part 3, Book 3 of the *Treatise* which focuses on natural virtues, compared to 124 appearances of obligation and 50 of duty in part 2 which focuses on artificial virtues. Data collected from *Hume's Texts Online* (URL = davidhume.org).

⁷ The two sections on justice and allegiance, which are tagged as artificial virtues in the *Treatise*, take up 10 out of 15 total appearances of the term ‘obligation’ in this work.

⁸ For an argument against Baier's suggestion, see Cohon (2021, 623–25).

1.1 Features of Hume's Conception of Moral Obligation

natural psychological principles. Bernard Wand comments that for Hume, the sense of duty is “partly natural and partly artificial” (1956, 161–62). I agree with this remark, and thence comes the title of this thesis: Hume's theory of obligation examines a part of the moral world which is ‘between nature and artifice’, which develops out of principles in human nature but is in tension with such principles, a tension that Hume attempts to explain from social and historical perspectives.

The sentimental basis of moral obligation, the relationship between virtue and duty, and the intersection of the distinction between virtue and duty, on the one hand, and that between natural and artificial virtues on the other, are the three main topics I will explore in this thesis. These topics do not exhaust what is important and/or interesting in Hume's relevant accounts. Noteworthy, I will *not* dive deeply into the heated debate on whether for Hume, the sentiment of moral obligation can motivate an agent to perform the obligatory action by itself, although for many scholars, this is the question which matters most for Hume's theory of obligation (I will return to this interpretative approach shortly). I will not do so because I think it would be more beneficial to deal with this question under the topic of whether moral sentiment, in general, can motivate on its own, and the latter topic involves a wide range of controversies on Hume's theories of passion, action, and moral sentiment the discussion of which would be distracting for the current research.⁹ In general, I align with the position defended by Elizabeth Radcliffe (2018, 125–33) that Humean moral sentiment *can* produce motives on its own, but I will stay neutral with respect to this controversy in the following discussions. The second topic I shall omit in this thesis concerns the relationship between moral obligation, especially the moral obligation to observe the three ‘fundamental laws of nature’, and the obligation imposed by positive laws. Although a detailed treatment of this topic is indispensable for a full-length study of Hume's theory of obligation, I could not produce it in the limited time and space permitted for this thesis. As Hume relies on his conclusions regarding justice to defend his position on political allegiance and the rights of civil magistrates, I take my work here as a necessary preparation for an investigation into the latter area.

This preliminary sketch only addresses some basic points by which Hume's conception of moral obligation can be understood. It does not touch on the questions of why his conception takes the shape it does and why I take these themes as central to his theory of obligation. The best way to answer these questions is to discuss what Hume's conception of moral obligation

⁹ For different views, see Brown (1988), Radcliffe (1996), Cohon (2010), Radcliffe (2018, chap. 5).

1.2 Obligation and Motivation

is *not*. I will make two negative points in the remaining sections of this chapter: first, Hume does not mean by obligation *as such* a determining motive, and the ‘necessitating’ force of obligation is not his primary concern in this part of his moral theory; second, Hume does not make obligation a dominant concept in his moral theory out of a group of considerations. These negative points help to identify what is special about Hume’s conception of obligation and what makes his conception interesting to study from a historical perspective. I leave the third negative point, namely, for Hume, moral obligation cannot simply be regarded as a form of approval of virtue or disapproval of vice, for Chapters 2 and 3, and I examine how Hume explicates and explains the difference between moral obligation and the evaluation of virtue and vice, through his distinction between natural and artificial virtues, in Chapters 4, 5, and 6. All these chapters will work together to give a full picture of the three features I listed above.

1.2 Obligation and Motivation

As I observed above, Hume does not provide any official definition for obligation as such or different types of obligations about which he talks in his arguments. This rather ‘careless’ style of writing makes it difficult for interpreters to presume a common starting point when they set out to talk about ‘Hume’s theory of obligation’. For many scholars, the central theme of such a theory is *moral motivation*: whether and how moral considerations can move one into action. This is the primary concern of Wand (1956, 155, 160–62), David Wiggins (1991), and Charlotte Brown (1994, 22–23, 28–31) when they discuss Hume’s account of obligation. Stephen Buckle (1991, 282) and Schneewind (1998, 371) maintain that for Hume, obligation as such is a synonym for “a determining motive”, and natural and moral obligations denote self-interested and moral types of motives.¹⁰ As far as I know, this is the most widely adopted perspective from which Hume’s account of obligation is investigated.

Though I choose to marginalize the theme of motivation in this thesis, I do not deny that it is an indispensable aspect of Hume’s account of obligation. The popularity of this perspective, moreover, has its justification. It makes up a part of the common understanding of obligation: to lie under an obligation to perform some action indicates that one is ‘bound’ to do so, which constitutes a reason for action which takes priority above others. The *Oxford English Dictionary* explains ‘obligation’ as meaning either (i) “moral or legal constraint; the condition of being morally or legally bound; the constraining power of a law, duty, contract, or (more

¹⁰ Also see Buckle and Castiglione (1991, 474), Cohon (2008b, 173), and Finnis (2011, 36–48).

generally custom, habit, etc.)” or (ii) “an act or course of action to which a person is morally or legally bound; what one is bound to do; a duty, commitment.”¹¹ In *A Dictionary of the English Language*, first published in 1755, Samuel Johnson explains ‘obligation’ as “the binding power of any oath, vow, duty; contract”¹², which corresponds to one aspect of the first definition in OED. To account for such a ‘binding power’, it appears, a theory of obligation needs to dive into human action and its motivation to investigate how it is possible for them to be bound and, more specifically, for a theory of moral obligation, how they can be bound by moral considerations.

This is also what Hume says when he first talks about duty and obligation in the *Treatise*, in his arguments against moral rationalists like Samuel Clarke who hold that moral values are determined by some “eternal fitnesses and unfitnesses or things, which are the same to every rational being that considers them” (T 3.1.1.4; 456). Hume starts his argument by stressing that morality is “suppos’d to influence our passions and actions”, and he supports this point with a piece of “common experience”:

[M]en are often govern’d by their duties, and are deter’d from some actions by the opinion of injustice, and impell’d to others by that of obligation. (T 3.1.1.5; 457)

Because morality influences people’s actions and affections, and because “reason alone ... can never have any such influence”, “[t]he rules of morality ... are not conclusions of our reason” (T 3.1.1.6; 457). Call this Hume’s ‘Motivation Argument’. Some scholars, like Buckle (1991, 272–83) and John Finnis (2011, 37–41), regard this argument as addressing Hume’s primary concern in positing a sentimentalist account of obligation: this account is intended to be a response to the rationalists, and what is characteristic about Hume’s view is that he identifies moral obligation with moral motivation or makes the latter a constituent aspect of the former. They also interpret Hume’s passage on the ‘is/ought’ distinction, which is presented as a side remark attached to his arguments against the rationalists, in the same light. The gap that Hume finds between ‘is’ and ‘ought’ statements, Buckle maintains, is not between fact and value, but between “obligations or motivations (and therefore actions) on the one hand, and facts and values on the other”; we cannot infer from ‘is’ to ‘ought’ because we cannot infer from truth statements to conclusions that can motivate us into action.

¹¹ ‘Obligation (n.)’, in *Oxford English Dictionary*, June 2024, <https://doi.org/10.1093/OED/9571286399>.

¹² ‘Obligation (n.s.)’, in Johnson (1755). Johnson notes two other meanings of this term: “an act which binds any man to some performance” and “favour by which one is bound to gratitude”.

1.2 Obligation and Motivation

There are, nevertheless, reasons to be sceptical about this interpretative approach. First, Hume's observation of the common experience about obligation and the Motivation Argument do not convey any determinate view on exactly *how* moral obligation motivates. This argument is open to several optional interpretations because for it to be valid, it is only required that moral obligation can participate in motivation in a manner that 'reason alone' cannot, and this requirement can be met no matter whether it motivates directly or indirectly, by itself or dependently on other motives or inclinations.¹³ Second, as long as it is possible to distinguish between moral obligation and other types of moral considerations like the approbation of virtue, as I suggested above (though I have not yet defended this suggestion), it is possible that Hume only talks about duty and obligation as a *case* for moral motivation in this argument. In other words, the Motivation Argument may only raise a more general point that moral judgements can participate in motivation in a way that 'reason alone' cannot; it does not say that all moral judgements *impose an obligation*, and it gives no indication as to whether 'the opinion of obligation' motivates in a manner like all other moral opinions or in a particular manner (for instance, by having a 'binding force' or authority to constrain other motives). In short, merely considering the Motivation Argument, Hume is not committed to either that (i) moral obligation motivates by itself or that (ii) all moral motivation takes the form of moral obligation.

One should indeed be careful about Hume's words on duty and obligation in this very first section of Book 3 of the *Treatise* because of this section's polemical function. Hume frequently talks about this topic in this section: he uses the duty against parricide as a representative case for the content of moral judgement (T 3.1.1.24-25; 466-68), and he talks about the influence of moral rules on the will by saying that they are considered as "universally forcible and obligatory" (T 3.1.1.23; 466). These remarks, however, cannot be taken as indicating that Hume himself takes duty and obligation as central or fundamental to moral evaluation and motivation, as it is arguable that in these arguments, Hume is attempting to defeat his opponents by using *their* language and proving that their theories cannot achieve their goals. Hume's target in Book 3, part 1 of the *Treatise* is to argue that "moral distinctions" are not "deriv'd from reason" but "deriv'd from a moral sense", as suggested by the titles of the two sections in this part. What Hume does not declare in his titles is that when shifting from the negative part of his arguments to the positive part, he also changes the primary reference of the phrase 'moral distinctions': in

¹³ Concerning the interpretation of the Motivation Argument, scholars are divided on both how to understand what Hume means by 'morality' and how to interpret his statement, previously argued in *Treatise* 2.3.3, that 'reason alone' cannot motivate. For an analysis of this argument and its probable interpretations, see Radcliffe (2018, chaps 4, 5).

1.2 Obligation and Motivation

Treatise 3.1.1, he talks about various types of moral concepts: right and wrong, virtue and vice, as well as duty and obligation; in *Treatise* 3.1.2, meanwhile, he confines his positive claims to virtue and vice (T 3.1.2.1-3; 470-71). It is risky, therefore, to regard Hume's anti-rationalist arguments as underscoring what is characteristic of his own theory of obligation. It is the *rationalists* whom Hume pushes to prove that 'the measures of right and wrong' have a 'necessary connection' to human will so that such measures could be "eternal laws, *obligatory* on every rational mind" (T 3.1.1.22; 465) because it is *they* who suggest that the "measures of right and wrong impose an obligation, not only on human creatures, but also on the deity himself" (T 3.1.1.4; 456). If, upon reading Hume's anti-rationalist arguments, one expects him to provide a theory in which moral judgements necessarily determine human will, one will be disappointed by what turns out to be Hume's positive account of obligation.

The Motivation Argument, at the same time, sheds light on Hume's theory of obligation in a crucial negative aspect. As Schneewind (1998, 370) notes, Clarke could respond to Hume's challenge by appealing to his theory of freedom and necessity, which Hume does not embrace. For Clarke, the fitnesses and unfitnesses of things impose an obligation on both God and human beings by constraining their free wills with a special type of necessity: *moral necessity*, which should be distinguished from physical necessity in efficient causation. God is under the obligation to act in conformity to such fitness relations though he has absolute freedom of not being determined by external causes; he is necessitated to do so by his perfect wisdom and benevolence. We, as rational creatures made in God's likeness, lie under the same necessity to conform our will to the fitnesses of the things. Hume, on the other side, argues against the distinction between physical and moral necessities; there is but one type of necessity, the idea of which arises from one's observation of the constant conjunction between objects, as Hume argues in his theory of causation (T 2.3.1.2-4, 2.3.2.1-4; 399-401, 407-410). Correspondingly, there is only one basis on which one can intelligibly claim that moral judgements 'necessarily' determine one's will, that is, by empirical observation of the regularity of their connections, and the rationalists are wrong in claiming that it can be proved *a priori* that moral rules are "forcible and obligatory" (T 3.3.1.23; 466). For Hume, the motivation of human actions is not different from the causal relations in the physical world (T 2.3.1.12-14; 403-04); we are not members of a kingdom of freedom who are exempted from natural causation but obliged by moral rules. Hume, moreover, it may be said, lacks a strong interest in developing a theory of practical reasoning from an agential perspective. In his study of human motivation and action, he more frequently takes up the perspective of a spectator, observing and explaining the causal

1.2 Obligation and Motivation

chains between mental states and actions as well as factors which influence this process. Motives, for him, are psychological states which can possibly cause voluntary actions, the scope of which, as he argues, is restricted to a set of passions.¹⁴ Although he also, at times, uses ‘reason’ in the sense of reason for action and even juxtaposes ‘reason’ with ‘motive’ (such as in T 3.2.1.9; 479), more frequently in his accounts of motivation and action, ‘reason’ only means demonstrative and probable reasoning which cannot motivate on its own (T 2.3.2.1-3; 413-14). It is Hume’s naturalistic approach to studying human motivation and action and his rejection of a distinctive type of moral necessity which underwrites his Motivation Argument against the rationalists.

From the perspective that Hume takes in his theory of action, then, the ‘binding force’ of moral obligation can only be conceived on the same footing as, and in competition with, the efficacy of non-moral motivating desires. The right question to ask about this ‘binding force’ is no longer whether in his theory, sentiment-based moral judgements can motivate *in principle*, but whether, based on empirical observation, moral perceptions can have sufficient *efficacy* to constrain other motives and make one feel ‘bound’. Although, in my reading, the sentiment of moral obligation, *qua* moral sentiment, can motivate by itself just like other types of moral sentiments can, it is usually too weak to counter-balance other motives. For it to ‘bind’ one into action, this sentiment needs the help of external factors like one’s desire for a good character and fear of bad fame, and the connection between the latter passions and moral sentiments depends on the education of parents and politicians. The binding force of the sense of duty, then, is the product of education, but the qualification of the sentiment of moral obligation to be strengthened in an ‘artificial’ manner is granted by its status as a ‘natural’ moral sentiment (T 3.2.2.24-27; 499-501). This position is consistent with the Motivation Argument, but it does not give the sense of duty any *natural* or *internal* authority to govern other motives.¹⁵ As Harris (2012) observes, Humean moral sentiments’ lack of authority constitutes one important reason why Hume’s contemporary readers, such as Henry Home or Lord Kames, Adam Smith, and Thomas Reid, were dissatisfied with Hume’s moral theory. Kames, for instance, criticizes Hume by arguing that in Hume’s system, “the noted terms of *duty*, *obligation*, *ought*, and *should*, &c, have no meaning.” This is not because Hume does not give these terms a place in

¹⁴ For Hume’s identification of motive and motivating passion, see *Treatise* 2.3.3 (413-418), 3.2.5.6 (518), 3.2.6.9 (532).

¹⁵ Hume does hold that moral sentiments have an “authority over our reason” and “command our judgement and opinion” (T 3.3.1.18; 584). This authority, however, influences our taste and imagination more strongly than our motives and actions (T 3.3.1.23; 586-87), and it needs the help of other motives to become an active principle.

his moral theory, but because Humean moral approval and disapproval are “too faint a principle to control our irregular appetites and passions” (*Principles of Morality*, 29-30). For Kames, an action is considered obligatory only when concerning it, we “have the consciousness of necessity, and of being bound and tied to performance, as if under some external compulsion” (*Principles of Morality*, 31), and this suggests that the sense of duty is a sentiment distinct from the approval of virtue and the disapproval of vice. Rather than making up a characteristic feature of Hume’s account of obligation, moral motivation is a problem hanging over his account.

Hume’s naturalistic view of motivation and action makes it attractive to relate his account of obligation to that of Thomas Hobbes, which makes up another reason to read his conception of obligation as falling under the category of motive or – which is applicable to both authors – motivating passions.¹⁶ This relation is especially noteworthy for interpreting another type of obligation apart from moral obligation, that is, ‘natural obligation’ which Hume explains as a sense of interest (T 3.2.2.23; 498). Knud Haakonssen, for instance, suggests that natural obligation for Hume is “closely akin to Hobbes’s idea” (1978, 15). Hobbes holds that “obligation begins where liberty ends” (*De Cive*, 2.10; 36); as liberty for him consists in “the absence of obstacles of motion” (*De Cive*, 9.9; 111), obligation, as the negation of liberty, consists in the obstacles which prevent agents from acting otherwise. In his *De Cive*, Hobbes distinguishes obligations into natural obligation and obligation which arises from agreements, and he further divides natural obligation into the type “where liberty is excluded by physical obstacles” and the type “where liberty is excluded by hope and fear; as when we say that a weaker man cannot disobey a stronger man whom he has no hope of being able to resist” (*De Cive*, 15.7; 174-75). The second type is especially important for our present concern: obligation here stands for a specific type of motivating passions; they are called ‘obligation’ because they prevent other passions from determining the will in a competition between motivating force and therefore restrict one’s liberty. Some of Hume’s remarks on obligation do bear a noteworthy affinity with this Hobbesian conception. In *Treatise* 2.1.10, for example, Hume talks about obligation in terms of “very strong motives of interest or safety to forbear” an action (T 2.1.10.6; 312) or some pleasure (T 2.1.10.9; 314), and he discusses such motives alongside external obstacles of actions, arguing that both factors influence our imagination about power and liberty. What makes this remark more intriguing is that Hume makes it in a section titled “of property and riches” where he gives his definition of property (T 2.1.10.1; 309-10), though he does not

¹⁶ For Hobbes’s conception of will and motivation, see especially *Leviathan*, 6.1-7, 6.50-54; 27-29, 33-34.

1.2 Obligation and Motivation

relate his remark on obligation and the imaginary idea of liberty to the definition of property. As Hume's account of obligation in Book 3 of the *Treatise* is closely intertwined with his theory of property, this textual connection is worthy of special attention.

Despite these marks of similarity, nevertheless, it is far from evident that Hume might have directly adopted a Hobbesian conception of obligation *or* natural obligation. Three reasons may discourage us from endorsing such an interpretation. First, Hume does not use the phrase 'natural obligation' in the passages cited above; he reserves this phrase for his theory of artificial virtues. Second, the Hobbesian conception cannot help to explain Hume's suggestion that the idea of obligation depends on conventions. Finally, this interpretation has difficulty with explaining the relationship between *natural* (or interested) and *moral* obligations unless one embraces the reading that moral obligation for Hume stays for motivating passions which are constituted or produced by moral sentiments, a reading which, though being popular, I have shed doubt on previously in this section. Hobbes does not maintain a distinction between natural and moral obligations. On a strong naturalistic reading of Hobbes, the obligations which have a moral bearing (such as the obligation to observe natural laws, including that to abide by agreements) can be reduced to psychological obstacles to liberty in one's deliberation for the end of self-preservation, and thereby be reduced to natural obligation.¹⁷ For Hume, meanwhile, the sentimental basis of moral judgements makes moral values irreducible to non-moral ones. The distinction and relationship between natural and moral obligations make up a central topic in Hume's account of obligation to which I will return in the next chapter and Chapter 6.

To sum up, while it is true that for Hume, both natural and moral obligations can motivate, there is, in my reading, no determinate textual ground to support the interpretation that these obligations just *mean* interested and moral motives. In the next section, I will discuss another perspective to delineate the characteristics of Hume's account of obligation which I find more promising as a starting point: the perspective which considers moral obligation first in terms of a type of *evaluative attitude* which arises from a spectatorial viewpoint and is then directed to an agent's own character traits and actions and influence their practical deliberation. The key question to ask from this perspective is not how moral obligation relates to motivation, but what type of moral evaluation it is, how it relates to other types of moral evaluations, and how, based on Hume's theory of human nature, this type of moral evaluation arises in human minds. These questions must be answered before one can proceed to discuss the motivational force of

¹⁷ For Hobbes's account of agreement-based obligation, see especially *De Cive*, 2.2-10; 34-36, *Leviathan* 14.6-7; 81.

moral obligation, and different answers to them have a significant influence on the latter topic. This is why I focus on the former questions in this thesis, and I will now turn to a central topic within this subject, namely, the relationship between obligation and virtue.

1.3 Obligation and Virtue

Apart from taking obligation as a synonym for a determining motive, another popular approach for delineating Hume's conception of moral obligation is to start from its nature as a sentimental response to character traits and actions and to question its relation to motivation or suspend this relation. Stephen Darwall (1995, 296) holds that moral obligation, for Hume, is the same as the sentiment of approbation of virtues. Other scholars, like Robert Shaver (1992, 547), Don Garrett (2014, 267), and Cohon (2021, 624), draw on Hume's words that "when the neglect, or non-performance of it [any action, or quality of the mind], displeases us *after a like manner*, we say that we lie under an obligation to perform it" (T 3.2.5.4; 517) to argue that moral obligation is a sentiment of disapprobation of the lack of certain character traits or omission of certain actions. I will argue for the second reading and discuss why the difference between these readings is important considering Hume's theory of moral sentiments in Chapter 3. For this introductory chapter, let me focus on a larger picture regarding the centrality of virtue in Hume's moral philosophy compared to the concepts of duty and obligation. I will start with a general observation of this centrality and then consider it from three perspectives: Hume's attitudes toward ancient and modern moral theories, his approach to moral studies and theory of moral sentiments, and his view on the relationship between morality and human nature.

As I discussed in Section 1.1, the central concepts in Hume's moral theory are virtue and vice. In Book 3 of the *Treatise*, after arguing against the rationalists, he immediately moves on to claim that "since *vice* and *virtue* are not discoverable merely by reason, or the comparison of ideas, it must be by means of some impression or sentiment they occasion" (T 3.1.2.1; 470, my italics) and that "[a]n action, or sentiment, or character is *virtuous* or *vicious*; why? because its view causes a pleasure or uneasiness of a particular kind" (T 3.1.2.3; 471, my italics). Also importantly, later in the *Treatise*, Hume claims that "[a]ctions themselves, not proceeding from any constant principle . . . are never consider'd in morality" as "[i]f any *action* be either virtuous or vicious, 'tis only as a sign of some quality or character" (T 3.3.1.4; 575). It is people's traits of character that the moral sentiments primarily respond to, and based on our sentiments, we ascribe the qualities of virtuousness or viciousness to these traits, and actions are only evaluated

1.3 Obligation and Virtue

derivatively as expressing one's character. In the second *Enquiry*, likewise, Hume sets the goal of this work by writing that:

We shall analyse that complication of mental qualities, which form what, in common life, we call Personal Merit: we shall consider every attribute of the mind, which renders a man an object either of esteem and affection, or of hatred and contempt; every habit or sentiment or faculty, which, if ascribed to any person, implies either praise or blame, and may enter into any panegyric or satire of his character and manners. (EPM 1.10; 173-74).

The language which occupies the central stage, in both works, is about merit and demerit, what is amiable and detestable, esteemed and contemptible. Though Hume also claims that “[t]he end of all moral speculations is to teach us our duty” (EPM 1.7; 172), he expresses no strong interest in presenting his moral inquiries as teachings on duties. He organizes Book 3 of the *Treatise* on the distinction between artificial and natural virtues, and the second *Enquiry* with his four categories of virtues: those useful to others, useful to oneself, agreeable to others, and agreeable to oneself. This is to be compared to another manner of organizing moral works in Hume's time, exemplified in Richard Allestree's *The Whole Duty of Man*, a popular devotional work which Hume read in his youth, and Samuel Pufendorf's *On the Duty of Man and Citizen* (the title of which was rendered as ‘The Whole Duty of Man’ in a seventeenth-century English translation), used as a textbook in Scottish universities, that is, the distinction between three categories of duties, those to God, to others, and to oneself (Allestree 1700; Pufendorf 2003). Although Hume maintains some distinctions between different types and aspects of duty and obligation, as I will discuss in Chapter 2, these distinctions are introduced to respond to questions which arise in his inquiries rather than to construct a system of duties.

As I discussed above, scholars like Baier consider the centrality of virtue in Hume's moral theory as expressing his preference for ancient moral systems compared to modern ones.¹⁸ One place where Hume expresses such an attitude is in his arguments against the distinction between so-called “natural abilities” and “moral virtues”, the latter including traits like benevolence and justice and the former, traits like “good sense and judgement”, “wit and humour”, and industry and perseverance (T 3.3.4.1-2, 3.3.4.7; 606-608, 610-11). One primary ground on which moral writers deny the moral merit of the latter traits is that they cannot be acquired voluntarily. This consideration cannot draw a reasonable distinction between these two groups of qualities since some “moral virtues”, especially “all the qualities which form the *great man*”, such as fortitude

¹⁸ For a similar view, also see Cohon (2008b, 162).

1.3 Obligation and Virtue

and magnanimity, are not voluntary (T 3.3.4.3; 608). People in their common life do not regard voluntariness as a condition of moral merit or demerit but “naturally praise or blame whatever pleases and displeases them”, and “the ancient moralists” did a better job of making their lists of virtues a loyal reflection of common sense (T 3.3.4.4; 609). In the second *Enquiry*, Hume draws on a group of ancient moral and historical texts to support his enlarged catalogue of virtues and expresses his favour for Aristotle’s *Ethics* and Cicero’s *On Duties* compared to *The Whole Duty of Man* in this respect (EPM App.4.10-18; 318-21). Why, then, do modern moralists emphasize the condition of voluntariness? Hume observes that:

In later times, philosophy of all kinds, especially ethics, have been more closely united with theology than ever they were observed to be among the heathens ... Philosophers, or rather divines under that disguise, treating all morals as on a like footing with civil laws, guarded by the sanctions or reward and punishment, were necessarily led to render this circumstance, of *voluntary* or *involuntary*, the foundation of their whole theory. (EPM App.4.20; 322)

Modern (which, in this context, pertains largely to ‘Christian’) moral writers draw an actually “imperceptible” distinction between virtues and merits because their sentiments and languages have been “warped from their natural course” (EPM App.4.20; 322) by a religion which, as Hume remarks in ‘A Dialogue’, “inspects our whole conduct, and prescribes a universal rule to our actions, to our words, to our very thoughts and inclinations, a rule so austere, as it is guarded by infinite, though distant, rewards and punishments” (EPM D.53; 342). Hume denies divine legislation as a necessary ground for moral values, and divine rewards and punishments as a necessary guarantee for virtuous conduct. Also, he argues against the merit of “monkish virtues” like “self-denial, humility, silence, solitude” which are only praised by sentiments perverted by “false religion” (EPM 9.3; 270) on the ground of “an extreme contempt of this life” compared to a future life (EPM D.55; 343). These views distance Hume’s position both from the theory of obligation maintained by natural law writers like Pufendorf, whose view I will discuss in Chapter 2, and the catalogue of virtues, centred on piety, humility, and a diligent self-inspection against pride, provided by Allestree;¹⁹ the same considerations may have partly motivated Hume to give a less prominent place to the conceptions of duty and obligation in his moral works.

Though Hume’s separation of morality from religion puts himself in opposition to ‘modern’ moralists, it is also to be questioned to what extent his moral theory embraces the spirit of the

¹⁹ For discussions of Hume’s response to Allestree’s *The Whole Duty of Man*, see, for instance, McIntyre (2014), Kroeker (2020).

ancients. Present-day philosophers who read Hume as a virtue ethicist disagree on whether, and in what sense, his virtue theory can be said as ‘Aristotelian’.²⁰ Compared to other philosophers in eighteenth-century Scotland like Francis Hutcheson and Reid, as Harris (2007, 233–35; 2010b, 5–6; 2011, 43–44) observes, Hume was more sceptical about the ancient idea that the aim of ethics is to show how to live a happy or prosperous life, to help with the cultivation of character, or to offer mental therapy. Hume’s view that virtues are approved for either their usefulness or their agreeableness, as Harris (2015, 135) reads it, implies his disagreement with the Stoic distinction of the *honestum*, or what is good in itself, from the *utile* and the *dulce*. This puts Hume in a position at odds with Cicero’s *On Duties*, a work composed under Stoic principles, in spite of the favour Hume expressed for this book, and as Aaron Garrett (2021, 193–95, 204–6) reads Hume’s references to Cicero in the second *Enquiry* concerning the diversity of virtues, they serve more as a tool to present his view in front of an audience who reads Cicero as in alignment with Christian moral theories, and to call for open-mindedness in philosophical debates in the spirit of Academic Scepticism. Compared to Aristotelians and the Stoics, Hume lacks theories of human *telos* and *eudaemonia*, of a harmonious cosmic order, or of reason as a ruling principle in human souls. Although he regards ancient moral theories as less distorted by religion and more in conformity with common sense, these remarks do not signify his subscription to the principles of ancient theories.

Apart from his secularized view of morality, Hume’s preference for the language of virtue and vice also has its ground in his study of human psychology, as a product of his “experimental method of reasoning into moral subjects”, as he declares in the subtitle of the *Treatise*, and of his ‘anatomist’ approach to moral inquiries, as he addresses at the end of this book (T 3.3.6.6; 621). As this analogy suggests, Hume took his project in this book as more of a description of our experiences with moral evaluation and an explanation of such experiences with his studies of human nature. Such a project is “subservient to *practical morality*” (T 3.3.6.6; 621) but does not take the direction of moral practice as its *primary* goal.²¹ From this perspective, virtue is a more fundamental moral concept because, first, this is how, according to Hume’s observation, our common-sense evaluative practice works (T 3.3.4.4; 609), and then, our moral sentiments, by their natural inclinations, primarily respond to people’s mental qualities which “extend over

²⁰ For different views, see, for instance, Greco (2013, 211–12, 217–20), and Swanton (2015, 56–61).

²¹ It is controversial whether Hume’s moral theory is purely descriptive and explanatory or, besides these functions, also normative. For different views, see Mackie (1980, 5–6), Schneewind (1990, 161), and Cohon (2008b, 161).

the whole conduct, and enter into the personal character” rather than single and transient actions (T 3.3.1.4-5; 575).

Why, then, does character take precedence over actions in moral evaluation? What plays a crucial role here is the close relation Hume observes between moral sentiments and ‘indirect passions’ like pride and humility, love and hatred, the examination of which takes up a major part of Book 2 of the *Treatise*. These passions are caused by things which can produce pleasure or pain by themselves and are closely related to (for instance, belonging to) a person, and they direct the mind’s attention to the idea of this person, making the latter such passions’ object (T 2.1.1-6; 285-87). Moral sentiments always produce one of the four indirect passions, which is “perhaps, the most considerable effect that virtue and vice have upon the human mind” (T 3.1.2.5; 473).²² This effect is “the most considerable” because the indirect passions play an indispensable role in shaping one’s *self-conception* in a social space: they function to establish a self-identity “as it regards our passions or the concern we take in ourselves” (T 1.4.6.5; 253) as they bind together one’s concern for various pleasant and distressing things to a self-concern, and, through the mechanisms of sympathy and comparison which connect one’s self-evaluation to one’s evaluation of others, these passions locate one in a place in a social community, a place of being loved or hated, esteemed or despised, the concern for which is of ultimate importance in practical life.²³ This “most considerable effect”, in my reading, makes moral evaluation more attracted to *durable* traits of character because they can be more closely related to the ideas of oneself or other persons and are, therefore, more suitable causes for indirect passions (T 2.1.6.7; 293), while single actions do not meet this criterion (T 3.1.5; 575). When one loves or hates another because of a single action, one still seeks the intention behind this action, and some durable quality expressed by this intention, to direct one’s passion to the idea of this person (T 2.2.3.4; 348-49). That moral sentiments primarily react to *persons* and *personal qualities* is an important feature of Hume’s theory, which may distinguish his version of moral sentimentalism from those of Hutcheson and Smith, for whom the moral sense or sentiments primarily react to people’s motives as token mental states.²⁴ This feature determines that the moral evaluation of

²² It is controversial whether Humean moral sentiments are by themselves types of pride or humility, love or hatred, or they are distinct perceptions which always produce such indirect passions. I agree with the latter interpretation, and I will come back to this debate in Chapter 3.

²³ For the indirect passions’ importance to the development of one’s self-conception, especially see the discussions of Baier (1991, chap. 6) and Ainslie (1999).

²⁴ See *Inquiry*, 2.2.1; 101; TMS 11.5-7; 18. Many scholars hold that Hume agrees with Hutcheson and Smith on this point. I will return to this controversy in Chapter 3.

1.3 Obligation and Virtue

actions is more *derivative* compared to the evaluation of characters, considering both the causal origin of moral judgements and the ground of moral values.

Hume's objection to modern authors' legalistic conception of morality, his descriptive and explanatory project in his moral studies, the primacy of characters over actions in accordance with the nature of moral sentiments, together with Hume's naturalistic understanding of human motivation and his lack of interest in developing a theory of practical reasoning, which I noted in the last section, all contribute to a *spectator-centred* approach to moral theory as suggested by Darwall (1995, 16, 286–87) and Donald Ainslie (2015, 291–95). Following this approach, the primary concern for a moral theory is to specify, as Hume suggests in the opening sections of both Book 3 of the *Treatise* and the second *Enquiry* as cited above, what we, as spectators to others' and our own characters, praise and blame, and to explain why we make such evaluations, rather than to ask 'what actions are right or wrong' or 'what I should do' from an agential perspective. This does not mean that the topics of duty and obligation have no place in such a theory but suggests that whatever place they may have, they need to be examined and explained in relation to the moral evaluation of characters in terms of virtue and vice.

Apart from Hume's attitudes toward ancient and modern moral philosophy and his theory of moral sentiments, another consideration which may contribute to the centrality of virtue in his moral philosophy is his view regarding the relationship between morality and human nature. As suggested by Michael Gill (2006, chap. 8) and Harris (2015, 51–77, 121–39), one central concern of Hume's moral inquiries is to respond to Bernard Mandeville and Hutcheson on the topic of whether morality arises from nature or human artifice, a question which dominates the investigation in Book 3 of the *Treatise* once Hume has established his sentimentalist position and starts to examine the principles of moral evaluation (T 3.1.2.6; 473). Mandeville argues in 'An Enquiry into the Origin of Moral Virtue' that human beings' natural inclinations are only concerned with their self-interest, and "the first Rudiments of Morality" are "broach'd by skilful Politicians, to render Men useful to each other as well as tractable" (*Fable*, 1:47). They do so by making use of people's pride, which "is so inseparable from his [human being's] very Essence" (*Fable*, 1:42): by instructing people to take pride in benefiting others, which actions are given the name 'virtue', politicians set up a reward to persuade people, especially the more high-spirited part of them, to deny their own nature and wage war against themselves to promote the public good. In his later work, which he included in his *Fable of the Bees* as 'Part II', Mandeville developed another account of the origin of morality not as a product of some politicians' design but as a result of the gradual development of human societies in history.

Hutcheson, on the other side, partly out of the aim to defeat Mandeville's earlier view, wrote his *An Inquiry into the Original of Our Ideas of Beauty and Virtue* to defend the position that both benevolent motives and the approval of them by the moral sense are natural to human constitution, can be felt independently of the influence of custom and education, and cannot be reduced to self-interested concerns.²⁵ As Gill and Harris observe, Hume took a middle ground between these two authors and, with this aim in mind, developed his distinction between natural and artificial virtues in the *Treatise*: he agrees with Hutcheson that human beings are not purely selfish and have both a benevolent instinct and a mechanism for moral approval by their very nature, but he agrees with Mandeville that this aspect of human nature does not suffice to enable them to live a sociable life, especially in large societies, and shares the view of later Mandeville that some moral institutions, especially 'justice' or what concerns the stability of property and the performance of promises, are the product of historical development.

How do these philosophers influence Hume's view of the relationship between virtue and obligation? Hutcheson's intention to defend the naturalness of morality motivated him to seek for the independence of morality from the conceptions of law and sanction, as such considerations give people self-interested reasons for beneficent actions. For him, the most fundamental moral concepts are moral goodness and badness, which are simple ideas perceived through the moral sense and occasioned by ideas of virtue and vice.²⁶ In the *Inquiry*, Hutcheson does not mention obligation until it comes to the final section, and he argues that obligation is a "complex moral idea" derived from the simple ones and defines moral obligation in terms of the moral sense's approval of some actions and disapproval of their omission.²⁷ In his *Illustrations on the Moral Sense*, Hutcheson criticizes other authors' use of words like 'duty', 'obligation', and 'ought' as "confused" and "unlucky" because they take obligation as an object of reason while it can only be properly understood as derived from the moral sense.²⁸ Some scholars hold that Hume took his definition of moral obligation and the division between natural and moral obligations from Hutcheson; though I think Hutcheson may have been an important source of influence, I do not read Hume as in complete agreement with his sentimentalist precursor in these aspects, which I shall discuss in more details in Chapters 2 and 3. What is noteworthy for my current concern is that Hume's agreement with Hutcheson on the point that there are *some* naturally virtuous principles also underlies his objection to the

²⁵ Especially see *Inquiry*, 2. Intro.; 85-88.

²⁶ See *Inquiry*, Preface, 1.1.9, 2. Intro.; 9, 23, 85.

²⁷ See *Inquiry*, 2,7.1; 177.

²⁸ See *Essay with Illustrations*, 146, 156.

1.3 Obligation and Virtue

legalistic conception of morality and his view on the derivative status of the idea of obligation. Importantly, for Hutcheson as well as for Hume, the existence of naturally virtuous qualities, especially benevolent motives, makes it possible that, at least in principle, one can act *virtuously* without a reflective awareness of this action's virtuousness, and one can fully discharge one's obligation in this domain even without any idea or sentiment of obligation. This is the first kind of moral duties, as Hume argues in "Of the Original Contract", "to which men are impelled by a natural instinct or immediate propensity, which operates on them, independent of all ideas of obligation ... Of this nature are, love of children, gratitude to benefactors, pity to the unfortunate" (E-OC 33, 479); duties which are representative examples for duties that correspond to 'natural virtues' in the *Treatise*. For both Hutcheson and Hume, this point is made against some natural law writers, like Pufendorf, whom I will discuss in the next chapter, that proper performance of duty depends on one's conscious obedience to the natural law and its obligation which *constrains* one's natural inclinations, and also against Mandeville, who finds all other-benefiting motives involving *self-denial*. Their positive views about human nature make moral obligation, especially, in the case of Hume, as how I read him, moral obligation as a painful, negative evaluation of one's actions and character, in some cases generating a passion of humility or self-hatred (T 3.2.1.8; 479), a *non-necessary* component for *a part of* our moral practice, although it is beneficial and, to a certain extent, indispensable since hardly anybody can be perfectly virtuous in such domains or always act as their virtuous principles would drive them to. Much more can be said about Hume's connection to Hutcheson in this respect, especially with regard to Hume's initial characterisation of 'natural virtues' in Book 3, part 2 of the *Treatise*, which I will analyse in Chapter 4.

As Hume's agreement with Hutcheson's positive view of a natural ground of morality may help to explain why obligation is not a central or fundamental concept in his moral theory, his *disagreement* with Hutcheson sheds light on those parts of his theory in which obligation plays a more central role than it does in their counterparts for Hutcheson. This is the second type of duty as Hume distinguishes them in 'Of the Original Contract', those which "are performed entirely from a sense of obligation, when we consider the necessities of human society, and the impossibility of supporting it, if these duties were neglected" (E-OC 34, 480), and this is the domain of what he calls 'artificial virtues' in the *Treatise*. Basically, the practice of artificial virtues like justice depends more strongly on the sentiment of moral obligation because Hume, unlike Hutcheson, does not regard our naturally altruistic motives and moral approval of them as sufficient for guaranteeing sociability when people are confronted with conflicts over

1.4 Chapter Preview

external possessions or difficulties in trusting strangers in transactions. The questions that Mandeville raises about the origin of virtue still work for this domain, and Hume agrees with Mandeville that *this part* of moral practice depends on human artifice both on the point that it underwent some gradual, historical development, and on the point that people's moral motive for justice in large societies, which is a strengthened sentiment of moral obligation, needs the help of the notions of honour and shame and the passions of pride and humility to be effective (T 3.2.2.25-27; 500-501), as I discussed in the last section. Hume's project to take a middle ground between Mandeville and Hutcheson on the relationship between morality and human nature, therefore, helps to explain the imbalanced weight he puts on obligation in his accounts of natural and artificial virtues.

Now let me sum up. In this section, I have gone through several reasons why obligation is *not* a central concept in Hume's moral theory: his critical attitude toward the legalistic conception of morality as well as the religion-dominated conception of virtue, which centres on a negation of this-worldly happiness, his spectator-centred approach to moral studies and the precedence of character compared to action in moral evaluation according to his theory of moral sentiments, and his agreement with Hutcheson on the naturalness of some virtuous qualities and our ability to approve of them in a disinterested manner. These considerations help to delineate the basic landscape of human beings' moral experience as observed by Hume, where people approve of a wide variety of useful or agreeable mental qualities as virtuous and disapprove of harmful or disagreeable ones as vicious. Against this background, moral obligation, taken as a form of moral evaluation, confronts Hume (and those readers following his steps) with a question: is there any difference between it and our approval of virtues and disapproval of vices? If there is, how to account for this difference? I will argue in later chapters that for this question, the distinction and connection between nature and artifice is Hume's strategy to give an answer: with his account of artificial virtues, he captures and explains what is distinctive about moral evaluation made in terms of obligation compared to those about virtues and vices. This will be the topic of my investigation for the remaining chapters, and I will now provide a brief outline for them.

1.4 Chapter Preview

This thesis has five main-body chapters. In Chapter 2, 'Natural Law and Hume's Taxonomy of Obligation', I discuss Hume's three distinctions between types of obligations: those between 'natural' and 'civil', 'natural' and 'moral', and 'natural' and 'artificial' obligations, all of which

are signs of the influence of early modern natural law theories on Hume's thought. I start by separating these different distinctions in terms of what is 'natural' or 'non-natural'. Then, I trace Hume's distinction between 'natural' and 'moral' obligations to Hutcheson's, and I argue that Hutcheson's distinction responds to two *aspects* of the obligation to observe the natural law as held by Cumberland and Pufendorf, namely, the moral value of an obligatory action and the force by which one is necessitated to perform it. Though the same does not apply to Hume's distinction, Hutcheson's works provide a historical clue for the separation of the binding power and moral evaluation related to obligation, and Hutcheson and Hume find different approaches for the unification of these two aspects. Finally, I compare Hume's natural/artificial distinction to Grotius's and Pufendorf's distinctions between connate and adventitious obligations on the one hand, and imperfect and perfect obligations on the other. I argue that against the general background of Hume's virtue-centred theory, a group of rights and obligations, i.e., perfect and adventitious ones, appear especially problematic and require explanation. Driven by this question rather than an aim to construct a system of obligations, Hume developed a theory of artificial virtues to accommodate these qualities in his theory.

In Chapter 3, 'Unpacking the Sentiment of Moral Obligation', I investigate Hume's delineation of moral obligation in the *Treatise* and discuss the causal origin, proper objects, and scope of this type of moral evaluation. I argue that for Hume, the sentiment of moral obligation is a displeasing sentiment about the absence of certain actions or mental qualities, and it differs from both the approbation of virtue and other forms of disapprobation both by its constitutive emotion and the ideas which typically cause it and are derived from it. This sentiment is most plausibly derived from the disapprobation of vice, either when one's failure to meet a certain expectation for virtue renders one's character morally deformed, or when the disapprobation of vice becomes associated with ideas of its remedy. Following the former route of generation, the objects of moral obligation are virtuous *character traits*. While some of these qualities are specified as one's disposition to be motivated by specific passions, like gratitude and humanity, others are not. In cases about justice or the 'three fundamental laws of nature', meanwhile, it appears that the objects of moral obligation are *actions*, which is a phenomenon that requires explaining. Finally, I examine two suggestions about conditions for some traits to be taken as duties, namely, whether they can be voluntarily acquired and whether they are required based on one's social roles and stations, and I argue that no sufficient textual evidence supports the idea that Hume endorses them as a standard for moral obligation.

In Chapter 4, ‘Moral Obligation and the Artificiality Arguments’, I examine the role moral obligation plays in Hume’s arguments for the difference between natural and artificial virtues, or ‘Artificiality Arguments’ as I call them. I argue that all four Artificiality Arguments in Book 3, part 2 of the *Treatise* are dominated by the same problematic, that is, the way in which ‘we’, or people who live and were educated in civilized societies, make moral judgements on rights and obligations about property and contracts in a way in tension with the natural inclination of our moral sentiments: such judgements are directed by universal and inflexible laws, they are applied to actions or objects in a strict and precise manner, and ‘we’ take the sense of duty as sufficient for characterising the virtue of justice and bestowing moral merit on just actions, without requiring a ‘natural’ virtuous motive to account for this moral merit. I argue that the Artificiality Arguments, especially the most controversial one in *Treatise* 3.2.1, rest on a supposition about a ‘state of nature’ concerning moral sentiments, a state in which people’s moral sentiments have not been influenced by conventions and education. This state is a tool for reasoning constructed by a method of separation, and the suppositions about such a state in *Treatise* 3.2.1 may have been motivated by Hume’s intention to respond to Hutcheson’s theory. With these arguments, Hume puts stress on the distinctive features of the moral obligation to justice and sets up the target for his explanation of the origin of justice.

In Chapter 5, ‘Sentiment and Convention: Hume’s Method for Artificial Morality’, I move on to discuss how Hume’s account of the *conventional origin* of justice can respond to the need to explain the ‘artificial’ characteristics of our *moral judgements* about justice, since it is more widely endorsed by scholars that these two topics are separated in Hume’s theory. I argue that the conventional origin of justice has a more substantial influence on moral judgements about justice by shaping people’s cognitive and evaluative habits, or customary mental associations, when they think in terms of justice and its rights and obligations. What is special about this habit is that we do not evaluate single actions of justice by their own consequences but by taking them as instances of a ‘general scheme’ of similar actions on a societal scale. The story of the origin of justice explains the development of a social institution as well as people’s mental habits developed within this institution, and Hume does so by offering a twofold or two-layered account of the history of justice: one of these layers, which I call ‘custom-convention’, describes a long causal chain extending from Hume’s suppositions about human nature to the practice of justice as we observe it now, and the other, which I call ‘agreement-convention’, is a philosophical reconstruction of the former process in terms of an agreement among rational individuals, which highlights the contribution of reason in the former process and characterises

the structure of the mental pattern, or the form of the special *general rules* which connect just actions to their consequences, that people follow, rationally or merely out of custom, to evaluate just actions.

Finally, in Chapter 6, 'From Interest to Morality: Convention and Obligation', I investigate how this historically originated and convention-dependent mental pattern may work to explain the 'artificial' characteristics of the rights and obligations of justice. I argue that this explanation works in two steps. First, the ideas of right and obligation gradually emerged as people obtained experience with conflicts over external possessions and coordinative settlements for them; the idea of obligation, at this stage, represented an 'artificial' interest to observe the property rules which constituted the 'natural obligation' of justice. These general ideas were associated by a special form of general rules which connected particular just actions to a societal scheme. Then, under the influence of *the same* general rules, people felt moral disapprobation of unjust actions by sympathizing with their victims who were conceived as participants of the same convention or as holders of property rights. This mental pattern is preserved or keeps being regenerated in the history of justice because of the underlying dynamics of human passions, imagination, and reasoning in social interactions, and it *partly*, together with the work of civil law and education, explains the way in which 'we' evaluate just and unjust actions in large, civilized societies. As following such convention-dependent general rules, just or unjust actions are characterised and evaluated in reference to the laws of justice rather than personal characters, we evaluate them differently compared to how we evaluate virtues and vices. Especially, when we form the *idea* of the moral obligation to justice, law and law-regulated action take precedence over character as the object of this obligation.

Chapter 7 is the conclusion of this whole research. Apart from summarizing the arguments I made in the main-body chapters, I make two points of observation based on this interpretation. First, our judgements about the obligation to have certain *natural* duties may instantiate a mixture of the ways in which we judge the obligation to justice and natural virtues and vices. Second, although different scholars have suggested that Hume's moral theory (or his theory of justice) should be read as a version of virtue ethics, rule-consequentialism, or contractarianism, his account of obligation, as I read it, does not support any of these readings fully. At the same time, Hume's account has aspects which assimilate these different moral theories; while it pursues a primarily descriptive and explanatory interest, it also sheds interesting critical lights on these normative theories.

Chapter 2

Natural Law and Hume's Taxonomy of Obligation

In the last chapter, I briefly delineated three features of Hume's conception of moral obligation: it is based on moral sentiments, it has a derivative status compared to the approbation of virtue and disapprobation of vice, and Hume discusses it mainly within a framework of the distinction between natural and artificial virtues. Then, I made two negative points about this topic: first, moral obligation as such does not imply a determinate motive, and second, Hume does not make obligation a dominant concept in his moral theory: his objection against legalistic conceptions of morality, his spectator-centred approach to moral inquiries, and his positive view of morality's ground in human nature, all lead him to give a central place to virtue rather than obligation. The question that arises from these discussions is whether the concepts of duty and obligation still have a distinctive place in Hume's moral theory, or whether they are simply different names for virtue and the moral approbation of it. My short answer is that the former is more plausible, and I will defend this interpretation in Chapters 3 and 4. Before that, as a continuation of my discussion of the intellectual context of Hume's moral theory, I would like to examine Hume's connection to another crucial source of influence on his conception of obligation, that is, early modern natural law theories, in this chapter.

The importance of the influence of natural law theorists, especially Grotius and Pufendorf, for the interpretation of Hume's moral theory, especially his theory of justice, has been stressed by a group of scholars such as Duncan Forbes (1975, chaps 1–2), Buckle (1991, chap. 5), and Haakonssen (1996, chap. 3). Hume defines justice as the observation of the “three fundamental laws of nature” (T 3.2.6.1; 526). Regarding his conception of obligation, one crucial aspect in which the obligation to justice differs from the moral approval of virtue and disapproval of vice is that it is “determined” by the *rules* of justice (T 3.2.6.6; 528), and that the idea of this obligation, as well as the ideas of ‘right’ and ‘property’, is dependent on the “laws of society; that is ... the laws of justice” to be intelligible (T 3.2.2.11; 490-91). The evaluations about virtue and vice, on the other side, are not regulated by rules or laws, especially not by laws so strict and universal as the laws of justice as to make judgements about them insensitive to particular situations and characters of the involved persons (T 3.2.6.9; 531-32). This aspect of

Hume's conception of obligation, as Harris (2010a, 25–33) suggests, can be interpreted in light of the influence of the Grotian tradition. For philosophers in the twentieth or twenty-first centuries, the topics of moral obligation and moral law may not seem very closely related; the same cannot be said of the seventeenth and eighteenth centuries, especially considering natural law theories' influence among intellectuals and in university education in Scotland in Hume's time.²⁹ For Hume, natural law theories not only represented an established conceptual framework with which he was engaged when developing his moral and political thoughts, but also, in various aspects, raised important questions and provided theoretical resources for him to make his response to these questions. Hume claims in a footnote in Appendix 3 of the second *Enquiry* that his theory of property and justice is “in the main, the same with that hinted at and adopted by Grotius” and quotes Grotius's *De Jure Belli ac Pacis* at length (EPM App.3.8n; 307). He appeals to Pufendorf to defend his theory of justice in his letter to Hutcheson (L 13; 1:33), and he also mentions Pufendorf's style of writing in a footnote about the state of nature in some editions of the second *Enquiry*³⁰, which proves his familiarity with these authors' works. Hume did not, like Hutcheson and Smith, in their positions as professors of moral philosophy, develop a systematic theory of natural jurisprudence. Both Hutcheson and Smith, as Haakonssen (1981, chap. 5; 1996, chap. 2) argues, attempted to draw a ‘synthesis’ between their moral sense or sentiment-based theories of virtue and vice, on the one hand, and natural law theories on the other hand. I think the same can be said of Hume's theory of justice. This provides a reason why the relationship, and especially the difference, between virtue and duty is important for Hume's moral theory.

Given the richness of early modern natural law theories of obligation which could have influenced Hume, especially given the profound differences between different authors' views, it would be imprudent for me to attempt to go through these theories in a temporal order here.³¹ Instead, I will take one specific question as my key thread in this chapter, namely, how natural law theories can help to disentangle Hume's taxonomy of obligation which sometimes appears perplexing to his readers. In the *Treatise*, Hume draws a group of distinctions between ‘natural’ and ‘non-natural’ obligations, and he does not always make it completely clear what he means by ‘natural obligation’ when using this phrase. Scholars like Buckle (1991, 63, 89–91), Darwall (1995, 287), and Schneewind (1998, 365–66) argue that Hume's distinctions were influenced

²⁹ For this aspect of Hume's intellectual context, especially see the discussions of Haakonssen (2012, 270–77).

³⁰ See Hume (1998, 218–19).

³¹ For very helpful discussions of these theories from which I benefit, especially see Forbes (1975, chap. 1), Buckle (1991, chaps 1–3), Haakonssen (1996, chap. 1), and Schneewind (1998, chaps 2–8).

by natural law authors. It would be helpful, therefore, to trace these probable sources and compare Hume's taxonomy to those of the natural law writers. I will consider the works of four authors in this chapter: Francisco Suárez, Grotius, Richard Cumberland, and Pufendorf. As Hutcheson provides an important point of reference about these authors' influence on Hume, I will also discuss him in this chapter.

In the following sections, I will start with Hume's clarification of the different meanings of the word 'natural' and separate his different groups of distinctions between 'natural' and 'non-natural' duties or obligations. By comparing the basic framework of Hume's distinctions and those of the natural law theorists, I argue that the most important difference between them is that Hume abstracts these distinctions from the conceptual framework which focuses on the relationship between legislators and their subjects and locates them in his causal explanation of the development of human societies and their moral notions. In Section 2.2, I will trace the change of the meanings of 'natural' and 'moral obligation', or their counterparts, from Suárez, Cumberland, Pufendorf, to Hutcheson and Hume. For the three natural law theorists, this pair of notions denote two *aspects* of the obligation to observe the natural law rather than two types of obligations. Hutcheson, out of the intention to criticize previous natural law theories, takes these two aspects apart into two different *meanings* of the term 'obligation' which respectively address the moral value of an obligatory action and the self-interested consideration by which one is necessitated to performing this action. Hume's distinction between 'natural' and 'moral' obligations may be an adaptation of Hutcheson's. While Hutcheson, in his later works, finds the unifying principle of these two types of obligations lying in divine legislation, Hume finds these types of obligations related by *causation*, namely, by the mechanism through which moral sentiments arise from people's conceptions of natural goods. Finally, in Section 2.3, I compare Hume's distinction between 'natural' and 'artificial' obligations with natural law distinctions between connate and adventitious obligations, on the one hand, and imperfect and perfect obligations on the other hand. I argue that Hume's distinction bears some apparent similarities to both these more traditional distinctions, but it does not match either of them. Rather, these distinctions raise questions about Hume's virtue-centred moral theory, and the natural/artificial distinction is Hume's response to them: he develops an account of the social and historical origin of justice to explain characteristics of our moral judgements about perfect adventitious rights and obligations which diverge from the natural inclination of our moral sentiments. This project occupies a central place in Book 3 of the *Treatise*. I finish this chapter with a comparison between the place of the distinction between natural and artificial virtues and obligations in the

Treatise and the second *Enquiry*, and I explain why it ceases to be a key topic in the latter work with changes in Hume's philosophical interest.

2.1 Hume's Taxonomy of Obligation

Hume does not provide any systematic taxonomy of different types of duties and obligations. He does not employ the widely accepted distinction between the duties to God, to others, and to oneself, which, as I noted in the last chapter, is adopted by Allestree's *The Whole Duty of Man*, Pufendorf's *On the Duty of Man and Citizen*, and Hutcheson's *A Short Introduction to Moral Philosophy*. He remarks that "[t]hat we owe a duty to ourselves is confessed even in the most vulgar system of morals" (EPM App.4.21; 322), and he uses this threefold distinction to argue against the criminality of suicide in his essay "Of Suicide"³², but these uses address more of an established conceptual framework than presenting his own taxonomy. Hume does not use the terms 'connate' and 'adventitious' obligations, or 'perfect' and 'imperfect' obligations, both of which, as I will discuss in Section 2.3, are maintained by natural law theorists like Pufendorf and accepted by Hume's contemporaries like Hutcheson. Instead of following such paradigms, Hume's distinctions between types of obligations, at least in the *Treatise*, are closely related to his use of the different meanings of the word 'natural', as he proposes a group of distinctions between 'natural' and 'non-natural' obligations understood in different senses. In this section, I will clarify these meanings of 'natural', tell apart different cases in which Hume uses 'natural obligation' with different meanings, and compare the framework of Hume's distinctions to those of his predecessors.

In Book 3, part 1 of the *Treatise*, after arguing that our 'moral distinctions' are derived not from reason but from sentiments, Hume immediately proceeds to investigate whether the principles of sentiment-based moral evaluation are 'natural' or not. Hume then distinguishes between five different meanings of the word 'natural', which will lead to different answers to this question. 'Natural' can be opposed to (i) what is miraculous, (ii) what is rare and unusual, (iii) what is artificial, and, as he adds in a footnote here, (iv) what is civil, and (v) what is moral (T 3.2.1.7-9; 473-74). In the first two senses, the sentiments of morality are definitely 'natural': they do not depend on miracles, and they can be very commonly observed in all human societies. In the third sense, in which 'artificial' means what depends on "the designs, and projects, and views of men", "it will appear afterwards, that our sense of some virtues is artificial, and that of others natural". Hume makes this claim together with a caveat, namely,

³² See E-Su; 580-88.

2.1 Hume's Taxonomy of Obligation

human designs and projects are but a part of nature: these principles are “as necessary in their operation as heat and cold, moist and dry” (T 3.2.1.9; 474). To suggest that our sense of some virtues is ‘non-natural’ in this sense, consequently, does not mean that such a sense does not *arise* from human nature. Neither does this suggest that the approbation of such virtues is uncommon or arbitrary. As Hume clarifies later regarding his statement that justice is an artificial virtue:

Mankind is an inventive species; and where an invention is obvious and absolutely necessary, it may as properly be said to be natural as any thing that proceeds immediately from original principles, without the intervention of thought or reflection. Tho’ the rules of justice be *artificial*, they are not arbitrary. Nor is the expression improper to call them *laws of nature*; if by *natural* we understand what is common to any species, or even if we confine it to mean what is inseparable from the species. (T 3.2.1.19; 484)

When opposed to ‘artificial’, then, ‘natural’ only denotes something’s independence from “the intervention of thought or reflection”. The distinction between ‘natural’ and ‘artificial’ virtues dominates Hume’s design of the structure of the remaining parts of this book: part 2 discusses artificial virtues, including justice, allegiance, and chastity; part 3 examines natural virtues like benevolence and prudence. This distinction, correspondingly, is the most prominent one in the *Treatise*, though it is much downplayed in the second *Enquiry*, to which I will turn later.

Unfortunately, Hume does not explain what he means by the term ‘natural’ when opposed to ‘civil’ or ‘moral’. What makes this lack of explanation more unfortunate for this thesis is that Hume applies both unclarified distinctions to duty and obligation. When he uses the phrase ‘natural obligation’, then, he means by it one out of three probable senses: ‘*non-civil*’ obligation, ‘*non-moral*’ obligation, or ‘*non-artificial*’ obligation; in some contexts, he does not make it immediately clear which meaning he has in mind. This raises requirements for both a careful separation of different distinctions when reading Hume’s texts and an investigation of probable historical sources of these distinctions, as one probable reason why Hume did not explain the natural/civil and natural/moral distinctions is that he presumed his readers’ acquaintance with these distinctions and took his distinctions as in conformity with previous ones.

Let me go through some samples of Hume’s use of all three meanings of ‘natural’ when talking about obligation. In his account of the origin of justice, Hume explicitly distinguishes between ‘natural’ and ‘moral’ obligations:

2.1 Hume's Taxonomy of Obligation

The *natural* obligation to justice, viz. interest, has been fully explain'd; but as to the *moral* obligation, or the sentiment of right and wrong, 'twill first be requisite to examine the natural virtues, before we can fully give a full and satisfactory account of it. (T 3.2.2.23; 498)

Hume makes it clear that 'natural obligation' pertains to interest, while 'moral obligation' is a moral sentiment, but he does *not* clarify what type of interest constitutes the natural obligation to justice or how the "sentiment of right and wrong" relate to other moral sentiments, like the approbation of virtue or the disapprobation of vice. Later, when discussing allegiance, he uses phrases like "the *moral* obligations of honour and conscience" and "the *natural* obligation of interest" (T 3.2.8.7; 545, also see T 3.2.9.3-4; 550-53), which refers to the same pair of concepts. Hume's delineation of obligation that "when the neglect, or non-performance of it [any action, or quality of the mind], displeases us *after a like manner*, we say that we lie under an obligation to perform it" clearly pertains to moral obligation, as it follows Hume's words that "all morality depends upon our sentiments", and "*a like manner*" refers to "*a certain manner*" in which we approve of virtues (T 3.2.5.4; 517). There is no direct evidence of how this delineation relates to natural obligation. I will refer to 'natural obligation' in this sense as 'interested obligation' in some places to avoid confusion with other meanings of the phrase 'natural obligation'.

One passage where Hume clearly uses 'natural obligation' in another sense, as pointed out by David Gauthier (1992, 405–9) and Don Garrett (2007, 262), occurs in an argument on the artificiality of promises:

A sense of duty supposes an antecedent obligation: And where an action is not requir'd by any natural passion, it cannot be requir'd by any *natural obligation*; since it may be omitted without proving any defect or imperfection in the mind and temper, and consequently without any vice. (T 3.2.5.6; 518, my italics)

'Natural obligation' here does not mean 'non-moral obligation', because Hume explains it with a piece of moral evaluation: if an action is required by a 'natural obligation', its omission will indicate a vice. In the same paragraph, Hume further illustrates this with the cases of the duties of humanity and parental care: "when we omit the duty, the immorality of the omission arises from its being a proof, that we want the natural sentiments" of humanity or parental solicitude (T 3.2.5.6; 518-19). Promissory obligation makes up an opposite example: it is required not because of any passion or motive in human nature, and this proves that fidelity to promises "is no natural virtue, and that promises have no force, antecedent to human conventions" (T 3.2.5.6; 519). I leave it for later chapters how this argument should be interpreted. For our present concern, it is sufficient to observe that this conclusion suggests that 'natural obligation' here

2.1 Hume's Taxonomy of Obligation

means 'moral' but 'non-artificial obligation'. It is arguable that Hume uses 'natural obligation' consistently in this sense in this entire section. Later, he declares that promises "impose no natural obligation" (T 3.2.5.12; 523), immediately after arguing that "interest is the *first* obligation" to performing promises (T 3.2.5.11; 523), which also suggests that 'natural' means 'non-artificial' here.

In another passage in the *Treatise*, the phrase 'natural obligation' means neither 'non-moral' nor 'non-artificial' obligation. This occurs in Hume's theory of political allegiance:

For as all government is plainly an invention of men, and the origin of most governments is known in history, 'tis necessary to mount higher, in order to find the source of our political duties, if we wou'd assert them to have any *natural* obligation of morality ... having found that *natural*, as well as *civil* justice, derives its origin from human conventions ... we shall find, that these two kinds of duty are exactly on the same footing, and have the same source both of their *first invention* and *moral obligations*. (T 3.2.8.4; 542-43)

It is clear that 'natural obligation' here does not mean 'non-moral obligation' since Hume calls it an "obligation of morality". Neither does it mean 'non-artificial obligation'. Hume uses this passage to argue against philosophers who reduce the duty to obey civil governments to that to keep one's promises, a duty which proceeds from one of the "three fundamental laws of nature" (T 3.2.8.4; 542). Both the duty of allegiance, which falls in the domain of "civil justice", and that of fidelity to promises, which belongs to "natural justice", in Hume's view, are established and draw their moral obligations from human conventions. In other words, both are 'artificial'. The 'natural obligation of morality' which could provide a source for political duties, therefore, should refer to a '*non-civil* obligation', or an obligation to observe 'natural', rather than 'civil', laws. Another case of Hume's use of the word 'natural' in the sense of 'non-civil' can be found in 'Of the Original Contract', where he remarks that "[t]he case is precisely the same with the political or civil duty of *allegiance*, as with the natural duties of justice and fidelity" (E-OC 35, 480). In his editorial note, Eugene Miller argues that Hume calls justice and fidelity 'natural duties' because he retreats from the distinction between natural and artificial virtues after publishing the *Treatise* (E-OC, 479). Miller fails to notice that the distinction between natural and civil duties is already present in the *Treatise* (also see T 3.2.8.5; 543), and calling justice a 'natural duty' does not need to indicate any change in Hume's view that it is 'artificial'.

It may be concluded from the discussion above that Hume draws three pairs of distinctions between types of obligations: 'natural' and 'moral', 'natural' and 'artificial', as well as 'natural' and 'civil' obligations. Some of these distinctions intersect with or are subordinated to another:

2.1 Hume's Taxonomy of Obligation

civil obligation, or the obligation to obey civil governments, is a subset of *artificial* obligation because it depends on conventions, and all *artificial* obligations, at least all those discussed in the *Treatise*, have *natural* (or interested) and *moral* subsets. Hume's use of the word 'natural', which is one of the most "ambiguous and equivocal" terms (T 3.1.2.7; 474), makes his taxonomy convoluted and sometimes perplexing. Based on the observation made above, it is possible to make his taxonomy clearer with the following chart:

Categories of Obligations	Natural (non-artificial)	Artificial			
		Natural (non-civil)		Civil	
		Natural (non-moral)	Moral	Natural (non-moral)	Moral
Examples	The moral obligation to take care of one's child	The interest in keeping promises	The moral obligation to keep promises	The interest in obeying governments	The moral obligation to obey governments

Once the structure of Hume's taxonomy of obligation has become clearer, a new question arises: why does Hume draw all distinctions between obligations in terms of what is 'natural' or 'non-natural'? One probable answer is that he just happened to have these three distinctions at hand as inherited from previous authors. Another probable answer is that his interest in the topic of obligation is closely intertwined with his concern with the relationship between human nature and morality, as I discussed in the last chapter, which makes him dedicated to separating more 'natural' parts or aspects of people's experience about duty and obligation from those less 'natural' parts or aspects. I think both answers may be true at the same time, as Hume's concern in the latter aspect may include an intention to engage with previous theories of obligation in which different 'natural/non-natural' distinctions are made. To see this, it will be helpful to first compare the basic frameworks of Hume's distinctions and those of his predecessors.

One reason why Hume's distinctions between natural and non-natural obligations appear perplexing is that it seems that these three distinctions are made on three different grounds. The distinction between 'natural' and 'civil' obligations, on the ground of the different *laws* from which these obligations are derived; the distinction between 'natural' and 'artificial' obligations, on the ground of their causal *origins*, namely, whether they depend on social conventions; and the distinction between 'natural' and 'moral' obligations, on the ground of their constituent

2.1 Hume's Taxonomy of Obligation

perceptions, namely, a sense of interest or a moral sentiment. It is difficult to tell how these distinctions may fit into one neat conceptual framework, or whether there is a conception of obligation *as such* which can unify all these aspects of distinctions. The similar distinctions in natural law theories, meanwhile, are made under one unified concept of obligation and one settled framework which focuses on laws and legislative acts. Let me say more about this.

A good starting point here is a definition of obligation which many early modern natural law theorists cite when discussing this concept: the definition in Justinian's *Institutes*, i.e., "an obligation is a legal bond, with which we are bound by a necessity of performing some act according to the laws of our State" (1913, 120).³³ The idea of 'legal bond' (*juris vinculum*) is central to those who theorized about obligation in the framework of natural law: obligation depends on *law* and imposes on the subject of this law an indispensable *necessity* to perform certain actions according to the law. Suárez, when tracing the etymology of the Latin word 'law' (*lex*), cites Thomas Aquinas who suggests that *lex* was derived from *ligandum* or 'binding'. 'Obligation' (*obligatio*) can also be traced to this origin (*ligare*); Suárez, accordingly, paraphrases Aquinas by arguing that "the true effect of law is to bind, or place under a binding obligation (*proprius effectus legis est ligare, seu obligare*)" (*Tractatus* 1.1.9; 24).³⁴ For Suárez, this 'true effect' characterises law in its strict sense: he defines law as "a common, just and stable precept, which has been sufficiently promulgated", and sufficient promulgation "implies an order for the purpose of creating an obligation" (*Tractatus* 1.12.4-5; 141-42). In a like spirit, Pufendorf defines law as "a decree whereby a superior obligates someone subject to himself to conform his actions to the superior's prescription" (DJNG 1.6.4; 120). For both authors, moreover, obligation is to be understood in the relationship between an 'obliger' or legislator, on the one hand, and those 'obliged' or subjected to an obliger's legislative act, on the other hand. For one to be bound by an obligation, as Pufendorf argues, one must live under a superior who has both the authority to rule and the strength to restrain one's internal freedom, and the latter condition distinguishes law from advice which has no necessitating force.³⁵

In this legalistic conceptual framework, it is easier to locate all distinctions between types or aspects of obligations. For instance, the distinction between 'natural' and 'civil' obligations, which Cumberland traces back to Justinian's *Digest* and which can also be found in Suárez and

³³ For quotes of this definition, see, for instance, DLN 5.10; 519; DJNG (1729) 1.6.5; 60; *Essay with Illustrations*, 146-47.

³⁴ For the Latin text, see Suárez (1679, 6).

³⁵ See DJNG 1.6.1, 1.6.7-8; 119, 122-23.

2.2 'Natural' and 'Moral' Obligations

Pufendorf, is made according to the types of laws from which these obligations are derived.³⁶ As Pufendorf explains it, natural obligation is imposed by the natural law prescribed by God, while civil obligation depends on laws prescribed and enforced by human authorities.³⁷ This is the predominant meaning of the phrase 'natural obligation' when natural law writers talk about it. Hume preserves the meaning of this distinction in a certain sense. For Hume, nevertheless, natural laws do not have a legislator, be it divine or human: they were established by human beings gradually through social coordination. Neither does the obligation to observe such laws depend on any relationship between the superior and the inferior: Hume does not talk about a significant inequality of power among the first inventors of such laws in the *Treatise*, and in the second *Enquiry*, he argues that a basic degree of equality in power must be found among those who participate in a scheme of justice (EPM 3.18; 190-91). Every convention member is *obliged* to observe conventional rules, but this arises from mutual pressure among them rather than a definite 'obliger' with superior strength. The relationship between law and obligation, correspondingly, has to be formulated in another way, which I will discuss in Chapter 6. As I shall not extend the scope of this thesis to Hume's theory of political obligation, I will not dive into more details concerning the relationship between natural and civil obligations here.

For the natural law theorists, the distinctions between 'natural' and 'moral' aspects of the obligation to observe the natural law, and those between connate and adventitious obligations and perfect and imperfect obligations, are also to be understood in light of their corresponding laws or aspects of legislative acts, which I will discuss in Sections 2.2 and 2.3. Regarding all these distinctions, the most drastic change between previous theories and Hume's is that Hume abstracts these distinctions from their original framework which focuses on the relationship between legislators and their subjects, and he locates them into a new framework of *causal explanation*, concerning how different notions of obligations could have emerged from human nature and developed into elements of moral institutions in civilized societies. Let me first turn to the natural/moral distinction.

2.2 'Natural' and 'Moral' Obligations

There is no determinate evidence concerning the source of Hume's distinction between 'natural' and 'moral' obligations. One probable source of influence, as suggested by Darwall (1995, 287),

³⁶ See *Tractatus*, 2.9.4; 253; DLN 5.10; 519.

³⁷ See DJNG (1729), 3.4.6; 256.

2.2 'Natural' and 'Moral' Obligations

is Hutcheson. Correspondingly, I will first take a look at Hutcheson's distinction and its probable sources and then turn to Hume's and compare these apparently similar distinctions.

Though, probably due to Hume's distinction, present-day scholars talk about Hutcheson's distinction as one between 'natural obligation' and 'moral obligation', it is noteworthy that Hutcheson himself did not employ the same phrasing: his distinction is between, on the one hand, 'natural good' and 'moral good', and on the other, two meanings of the term 'obligation'. Natural good, as Hutcheson explains it, is a synonym for 'interest' or 'advantage' which we pursue by self-love, while moral good is a quality of actions which we approve of with our moral sense in a disinterested manner.³⁸ The two meanings of the term 'obligation' echo this distinction. As Hutcheson renders it in his *Illustrations on the Moral Sense*:

When we say one is obliged to an Action, we either mean, 1. That the Action is necessary to obtain Happiness to the Agent, or to avoid Misery; Or, 2. That every Spectator, or he himself upon Reflection, must approve his Action, and disapprove of his omitting it, if he considers fully all its Circumstances. (*Essay with Illustrations*, 146)

If one takes this distinction as one between two species of obligations under one common genus, difficulties will arise concerning what this genus is. The former meaning addresses a necessity to perform an action, but the latter does not: in his explication of this distinction, Hutcheson makes it clear that if by 'obligation', one means an 'indispensable Necessity', this only agrees with the former meaning.³⁹ By such a necessity, obligation in its former meaning can work as a motive, but Hutcheson does not mention motivation in its latter meaning.⁴⁰ Moreover, these two ideas of obligation are derived from different faculties of the mind: the former from desires for self-interest, and the latter from the moral sense. Darwall (1995, 233–35), for this reason, remarks that there is a genuine puzzle in Hutcheson's taxonomy of obligation.

One way to get out of this difficulty is to read Hutcheson's distinction as a response to a long tradition of separating the 'natural' and 'moral' *aspects* of the natural law and its obligation

³⁸ See *Inquiry*, 2.Intro.; 85-86.

³⁹ See *Essay with Illustrations*, 146. Partly based on this reason, I am sceptical about Michael Walschots's (2022, 132) suggestion that the two meanings of obligation find their unity in a conception of necessity.

⁴⁰ Hutcheson mentions motivation in his definition of obligation in its moral sense in his *Inquiry*; see *Inquiry* 2.7.1; 177. He does not do so, however, in all his later works, including the *Illustrations*, *A Short Introduction to Moral Philosophy*, and *A System of Moral Philosophy*: the distinctions in the two latter books are more similar to those in the *Illustrations*. See *Short Introduction*, 2.2.2; 112-13; *System* 2.3.6; 264. I largely agree with Darwall's reading that the Hutchesonian moral sense cannot motivate by itself, and obligation in its moral sense, which is an idea derived from the moral sense, cannot by itself serve as a motive. Therefore, like Darwall (1995, 218–23), I give more authority to Hutcheson's definitions in his later works. For different interpretations of Hutcheson's definition of moral obligation, though, see Buckle (1991, 280–81), Jeffery Edwards (2006, 25), and Aaron Garrett (2007, 251).

2.2 ‘Natural’ and ‘Moral’ Obligations

rather than a distinction between two species of obligations. This consideration will also help with the interpretation of Hume, although important differences can be observed regarding the ‘natural/moral’ distinction both between the natural law theorists and Hutcheson and between Hutcheson and Hume. Hutcheson mentions a number of authors when making this distinction: Cumberland and Pufendorf in the *Inquiry*, and Jean Barbeyrac in the *Illustrations*⁴¹, which supports such a connection.

For the authors Hutcheson mentions in these texts, the distinction between what is ‘natural’ and ‘moral’ addresses a crucial question in Christian natural law theories, namely, whether God prescribes the natural law based on some antecedent criteria of good and evil, or good and evil are determined by God’s free decree. Christian moral theorists have reasons for worries on each side: taking an extreme position on the former side would make it possible for one to lead a just life without self-consciously conforming one’s actions to God’s legislative will, while an extreme position on the latter side would make God’s legislation arbitrary, so that it would be possible that actions such as murder could be rendered good. One paradigmatic approach to taking a middle ground between these extremes was suggested by Suárez with the distinction between the *indicative* and *preceptive* parts of the natural law: the former proceeds from God’s intellect and indicates what is good and evil, while the latter proceeds from God’s legislative will, imposes an obligation on free and intelligent creatures like human beings and directs their actions.⁴² Correspondingly, Suárez distinguishes between two phases of duty: one phase, which he calls “natural duty” (*debitum naturale*), consists of “an intrinsically fit proportion between the object and the act in relation to right reason, or rational nature”, while the other is a “moral obligation” (*obligatio moralis*) which “proceeds from the law itself, as an effect of the law” (*Tractatus* 2.9.4-6, 2.15.17-18; 253-55, 341-42). The denomination ‘moral’ here, as opposed to ‘natural’, then, underscores God’s legislative will and its effects.

Roughly the same may apply to Cumberland’s and Pufendorf’s distinction between natural and moral goods by which Hutcheson’s distinction could have been influenced. Both authors understand natural good, as related to human beings, as the quality of things which serves their preservation and perfection, and moral good as a quality of voluntary actions which consists in their conformity to law, especially the natural law.⁴³ Compared to Suárez, though, they take a more voluntarist position: while Suárez acknowledges some moral value intrinsic to actions in

⁴¹ See *Inquiry*, 2.7.1; 178; *Essay with Illustrations*, p. 146.

⁴² See *Tractatus*, 2.6.5-13; 210-21.

⁴³ See DLN 3.1, 5.9; 462-63, 515-16; DJNG 1.4.4, 1.7.3; 113, 129.

2.2 ‘Natural’ and ‘Moral’ Obligations

relation to human nature, which ground the ‘natural duty’ to perform them, Cumberland and Pufendorf make the separation between the ‘natural’ and ‘moral’ domains more absolute. For Pufendorf, especially, who structures the framework of his theory with the distinction between ‘natural’ and ‘moral entities’, all moral qualities are ‘superadded’ or ‘superimposed’ on created natural things by the wills of intelligent beings.⁴⁴ Therefore, there could be no such thing as a ‘natural duty’; duty, for Pufendorf, is just “human action in conformity with the commands of law on the ground of obligation” (*De Officio*, 1.1.1; 17).⁴⁵ For them, as it is for Suárez, the conformity between God’s intellect and legislative will guarantee the *conformity* between the natural law and what is fit or unfit for the preservation and perfection of human beings, and consequently, between what is morally and naturally good about human actions. Cumberland talks about their relationship in the language of *form* and *matter*: the matter of the natural law pertains to the “*Nature of Things*, especially of *Man*, and also of the *common Good*” (DLN Intro. 9, 256), while its form gives it the proper nature of a law; the combination of these two phases makes the natural law always promote the common good and the happiness of individual persons.⁴⁶ Pufendorf also argues that the natural law is adapted to human nature “in such a way that its observance is always linked with men’s utility and advantage” (DJNG 2.3.18; 154). This conformity, nevertheless, cannot be read as suggesting that the moral good can be reduced to or derived from the natural good, as they have different origins and are perceived, for us, in different ways: natural good by desires and instrumental reasoning, while moral good, by the right reason’s recognition of the natural law.⁴⁷ As Buckle puts it, for Pufendorf, utility “is therefore a feature of natural law, but not its ultimate foundation” (1991, 67).

The distinction between natural and moral goods has an important influence on these two authors’ accounts of obligation. For both, as for Suárez, obligation characterises law proper and makes up a part of a law’s ‘formal’ or ‘moral’ aspect. Cumberland defines obligation as “that Act of a Legislator, by which he declares, that Actions conformable to his Law are necessary to those, for whom the law is made” (DLN 5.27; 554); Pufendorf disagrees with this and holds that obligation is a moral quality inherent in the obliged person, “through which a person is

⁴⁴ See DJNG 1.1.2-5; 99-101.

⁴⁵ Compared to Pufendorf, the views of Grotius and Barbeyrac are closer to that of Suárez: Grotius suggests that actions “are in themselves either Obligatory or Unlawful (*debiti, aut illiciti per se*), and must, correspondingly, be understood to be either commanded or forbid by God himself”. Barbeyrac argues that Grotius is not entirely right and that the “indispensable Necessity, which is implied in the Idea of Obligation” can only come from the will of a superior, but he agrees with Grotius that there is a “natural Morality of Actions” which arises from the fitness or unfitness of things. See DJBP 1.10.2; 151-52. Noteworthy, Grotius’s phrasing here is *debiti* rather than *obligationes*, and Barbeyrac rendered it into ‘obligatoires’ in his French translation.

⁴⁶ See DLG Intro. 9-15, 5.1; 256-63; 495-96.

⁴⁷ See DJNG 2.3.13-14; 149-51.

2.2 'Natural' and 'Moral' Obligations

bound by a moral necessity to do, or to admit or suffer something” (DJNG 1.6.5; 121). For there to be a *necessity* which constitutes an obligation, though, for both authors, the legislator must attach to their law some *sanctions*, i.e., rewards and punishments which are naturally good or bad to the obliged party.⁴⁸ Cumberland distinguishes between “the First and Principal Cause” of the obligation to observe the natural law, that is, “the Will and Counsel of God” which provides obligation with its moral ground, and what motivates human beings to such observance, that is, rewards and punishments which consist primarily in the beneficial or harmful consequences of actions in accordance to the natural order of created things.⁴⁹ A proper awareness of one’s obligation must include both aspects and, in an order of discovery according to human nature, we usually start with the latter, or the ‘natural’ aspect, and proceed to realize the former, ‘moral’ aspect.⁵⁰

For Pufendorf, the idea of ‘moral necessity’ he uses in defining an obligation distinguishes it from physical necessity which deprives one of the freedom to choose otherwise. For there to be such a moral necessity, meanwhile, there must be “*external* principles strong enough to place a bridle on his internal freedom” (DJNG 1.6.8; 123, my italics) which proceed from the legislator. Pufendorf argues that a free agent’s will can be affected in two manners, either ‘naturally’ or ‘morally’:

[T]hough many other things may cause the will to incline in one of two directions, obligation nonetheless has this special advantage over them. While they press upon the will with a sort of natural weight after whose removal it returns of itself to its former indifference, obligation affects the will morally by imbuing it with a special internal sense that compels it to censure its own actions and deem itself blameworthy if it has not conformed itself to the prescribed norm. (DJNG 1.6.5; 121)

The interpretation of this passage is controversial. Schneewind (1998, 137–38) and Darwall (2012, 213–38) read it as indicating that obligation, unlike natural desires, only motivates morally, while Buckle (1991, 62–63) and Thomas Pink (2009, 108–14) read it as suggesting that obligation *both* affects the will with some ‘natural weight’ *and* carries with itself a moral awareness that acting otherwise would be blameworthy. I think the latter reading accords better with the text, as Pufendorf, immediately after this passage, argues that nothing can impose on the human mind “the necessity of either doing or omitting something” apart from “thoughts of

⁴⁸ Cumberland emphasizes that sanctions, in his theory, include both rewards and punishments. See DLN Intro. 14; 260-62. For Pufendorf, it only refers to punishments. See DJNG 6.14; 127.

⁴⁹ See DLN 5.22; 543-44. For helpful commentaries on this distinction, see Haakonssen (2000, 37–41).

⁵⁰ See DLN 5.45; 606-607.

2.2 'Natural' and 'Moral' Obligations

the good and evil that will arise for others or ourselves from what we do” (DJNG 1.6.5; 121). Also, when comparing obligation with coercion, he argues that “each of them ultimately points to some thing to be feared”, and what makes obligation different from coercion by bare natural force is that it also forces the obliged person to acknowledge that “the evil befalling him ... is not undeserved” (DJNG 1.6.5; 122). Schneewind puts much stress on Pufendorf’s claim that an obliged person reacts to their legislator, who has both the authority and strength to rule, with “a fear mingled with reverence”. Reverence, which arises “from a reflective consideration of the reason”, suffices to make the obliged one observe the legislator’s will “apart from fear” (DJNG 1.6.9; 123). Noteworthy, however, Pufendorf remarks on this rational observation that it happens “*as we accept advice*” (DJNG 1.6.9; 123, my italics), and advice does not give rise to an obligation by itself.⁵¹ While one can observe the natural law both by rational reflection and by a fear of deserved punishment, then, it seems that the latter is indispensable for there to be an obligation.

If my reading is correct, then, for both Cumberland and Pufendorf, the obligation to abide by the natural law has both a ‘moral’ and a ‘natural’ aspect: the former accounts for the moral value of obligatory actions, making the omission of such actions blameworthy, while the latter makes obligation ‘binding’ or makes it ‘necessary’ for the obliged person to observe the law on the basis of their fear for what is naturally evil (and, for Cumberland, desire for what is naturally good). This agrees with Hutcheson’s distinction between two meanings of the word ‘obligation’ in the *Illustrations*, the latter about moral evaluation and the former about the ‘necessity’ of an action to pursue natural goods and avoid natural evil. Later in the same paragraph, Hutcheson comments on Barbeyrac’s commentary on Grotius which suggests that obligation denotes “an indispensable necessity to act in a certain manner” that this means that a powerful superior makes obligatory actions necessary for obtaining happiness and avoiding misery.⁵² Also in the *Inquiry*, he appeals to Cumberland and Pufendorf to argue that the interested sense of obligation can be derived from the consideration that sociable actions are the best means to promote “the natural Good of every Individual” (*Inquiry*, 2.7.1; 178). It is possible, then, that Hutcheson was self-consciously responding to the divisions of different aspects of obligation of these natural law authors in his distinction rather than raising a distinction between two types of ‘necessitating’ motives.

⁵¹ DJNG 1.6.1; 119.

⁵² *Essay with Illustrations*, 146. For Grotius’s and Barbeyrac’s views on obligation, see Note 45.

2.2 'Natural' and 'Moral' Obligations

The most significant difference between Hutcheson's conception of obligation from those of Cumberland and Pufendorf is that he takes this distinction out of its original legalistic context. Hutcheson defines obligation in its moral and interested senses in the *Inquiry* to respond to the question of whether we can have a sense of obligation abstracting from the laws of a superior, and obligation defined in terms of law and sanctions is a third sense of this term.⁵³ Importantly, for Suárez, Cumberland, and Pufendorf, when 'moral' is opposed to 'natural', it refers to an aspect of the subject matter in relation to a law and a legislative act; for Hutcheson, meanwhile, 'moral' refers to what is perceived by the moral sense, and the moral sense approves of actions independently of ideas of a law or a lawgiver. This sentimentalist turn may have been motivated by two reasons. First, especially in the *Inquiry*, as suggested in the title of this work, Hutcheson was primarily concerned with the psychological origin of our aesthetic and moral ideas. The central question here is from what faculties of the mind we derive our ideas of moral goodness and obligation rather than what is these qualities' first and principal cause: this work tracks the order of discovery rather than that of creation. It is, then, one thing that moral obligation may ultimately be determined by the natural law prescribed by God but another thing that we first obtain its idea with our moral sense. Second, Hutcheson was dedicated to defending human nature's self-sufficiency in performing morally good actions (which, for him, are those actions performed out of benevolent motives) and approving of these actions' moral goodness.⁵⁴ He was keen to show that benevolent motives and the moral sense are original to human beings' natural constitution and cannot be reduced to other origins, especially self-interest, and he was dissatisfied with Cumberland's and Pufendorf's theories because moral evaluation, for them, relies too much on rational reflections which are "too slow, too full of doubts and hesitation ... to direct our Actions for the Good of the Whole" compared to the moral sense (*Inquiry*, 2.7.3; 180, also see 2.1.4; 94). Though reflections on interest and a superior's law have their place in Hutcheson's conception of obligation, consequently, it is important for him that we have access to the idea of obligation, taken in its 'moral' sense, independently of these considerations.⁵⁵

Once divine legislation fades out of the horizon, nevertheless, it immediately becomes a question where lies the principle of unity of the different meanings of obligation as Hutcheson explicates it. The different aspects of the obligation to observe the natural law, for Cumberland and Pufendorf, find their unity in God's legislative act, and the conformity between natural and

⁵³ *Inquiry*, 2.7.1; 177-79.

⁵⁴ Especially see *Inquiry*, Intro., 85-88.

⁵⁵ For Hutcheson's distinction between different meanings of the work 'obligation' as a response to Cumberland and Pufendorf, also see the discussion of Aaron Garrett (2017, 23-26).

2.2 'Natural' and 'Moral' Obligations

moral goods involved here is guaranteed by the conformity between God's intellect and will. Following Hutcheson's approach, meanwhile, there is no reason evident at first sight that what is morally obligatory for a person is also naturally obligatory or sanctioned by some law. In his later work, *A Short Introduction to Moral Philosophy*, where he combines his moral sense theory with a system of natural jurisprudence, Hutcheson argues that while we first obtain our ideas of right and obligation by our moral sense and independently of considerations of law or command, we can, once we have observed and reflected sufficiently on the nature of things as God's creation and "have ascended to the notion of a natural law", we can define these notions by referring to the natural law. The coherence between these two approaches for defining these notions is guaranteed by the point that "the grand aim of the law of nature is the general good of all, and of every part as far as the general interest allows it" (*Short Introduction*, 2.2.1; 111).⁵⁶ In other words, harmony can be presumed between different senses of obligation, according to Hutcheson's later theory, due to the benevolent intention of God in creating human nature (including human beings' naturally benevolent motives and their moral sense) and prescribing the natural law as it is in the world where we find ourselves.⁵⁷

Hume's distinction between 'natural' and 'moral' obligations has an apparent similarity with Hutcheson's: natural obligation consists of a sense of interest, and moral obligation refers to a "sentiment of right and wrong" (T 3.2.2.23; 498). Hume also shares Hutcheson's interest in investigating the psychological origin of the idea of obligation and using the word 'moral' for what is perceived through moral sentiments. Also, like in the case of Hutcheson's distinction, while natural obligation can directly serve as a motive, moral obligation, for Hume, primarily addresses a type of *moral evaluation* and does not bear with itself any *necessity* for action: as I argued in the last chapter, such a necessity can be attached to the sentiment of moral obligation artificially by education but is not part of its natural force. As Hume does not refer to the natural law theorists' accounts of obligation in his discussion, there is no evidence whether he had their distinctions of different aspects of obligation in concern, but there may be an indirect influence through Hutcheson's distinction.

It is also important to observe three differences between the 'natural/moral' distinctions of Hutcheson and Hume. First, Hume does not, as Hutcheson does, define natural obligation or obligation taken in a sense related to interest or 'natural good' as what is necessary, upon

⁵⁶ I omit the words added by the anonymous English translator of this work when citing it.

⁵⁷ It is controversial, though, whether Hutcheson's textbooks on natural jurisprudence were development of his own philosophical theories or were products of mere pedagogical consideration. For different views, see Moore (1991) and Haakonssen (1996, chap. 2).

2.2 'Natural' and 'Moral' Obligations

rational calculations, to obtain happiness and avoid misery. Despite the similarity between their distinctions, it cannot be presumed that Hume would endorse Hutcheson's definitions, and as I will discuss in Chapter 3, their depictions of moral obligation are not completely the same. Second, Hutcheson emphasizes that both the moral and natural meanings of 'obligation' can be conceived independently of the idea of a law. Interestingly, Hume's position at this point may be closer to the natural law theorists' when he suggests that the idea of obligation is unintelligible independently of the law of justice (T 3.2.2.11; 490-91), which suggestion can at least be applied to the obligation to observe the three 'laws of nature'. To recall, for Hume, the obligation to observe the natural law is not the obligation to obey the command of a divine or human legislator; both the so-called natural laws and the obligation to observe them depend on conventions. This common artificial origin of the law of justice and its obligations, about which Hutcheson was critical in his comments on Book 3 of the *Treatise*,⁵⁸ makes the ideas of law and obligation more closely related in Hume's picture. Third, and most importantly, as Hume excluded divine providence and legislation out of the ground of moral values more thoroughly than Hutcheson, natural and moral obligations, for him, cannot be unified by God's benevolent intention. How, then, do these two obligations relate to each other? Hume remarks that one way in which they are related is *causation*: the natural obligation to allegiance, for instance, causes its moral obligation (T 3.2.9.3; 551). Hutcheson would disagree with such a relation because for him, the moral sense, from which the idea of moral obligation arises, is an original faculty of the human mind and cannot be explained by the work of other faculties. For Hume, however, moral sentiments do not arise from an original mental faculty but from sympathy and ideas of what is useful or agreeable to people, or, in Hutcheson's term, ideas of what is naturally good.⁵⁹ This does not mean that moral values, for Hume, can be reduced to natural ones because the former, like in the theory of Hutcheson, depend on moral sentiments, but Hume's version of sentimentalism allows a causal relation between ideas of natural and moral goods in a way that Hutcheson's does not. If there is any principle of unity between natural and moral obligations in Hume's account, in my reading, it is to be found in this *causal* relation, which I will discuss in Chapter 6 of this thesis.

⁵⁸ See Hume's response to Hutcheson's critique in his letter to Hutcheson in September 1739 (L 13; 1:33).

⁵⁹ For a helpful discussion of the differences between Hutcheson's and Hume's theories in this aspect, especially see Gill (2006, chaps 14, 17).

2.3 'Natural' and 'Artificial' Obligations

What, then, about Hume's distinction between 'natural' and 'artificial' moral obligations, or the obligations which correspond to 'natural' and 'artificial' virtues? Two different suggestions have been made in the literature about its probable sources: Buckle (1991, 89) argues that this distinction "is, in outline, the same as Pufendorf's distinction between the natural and the adventitious"⁶⁰, while Schneewind (1998, 365–66) suggests that it reflects Hume's attempt to accommodate the distinction between imperfect and perfect duties in natural law theories in his moral theory. Though both suggestions can find some support in Hume's characterisations of the natural/artificial distinction, I am sceptical about both of them: no neat correspondence, in my view, can be found between Hume's natural/artificial distinction and such previous distinctions. On an apparent level, this is because the items in these categories do not cohere. On a deeper level, this is because the natural/artificial distinction, more than the distinctions discussed above, is closely related to Hume's problematics and methodology in the *Treatise* and reflects what is unique and original to his theory. To see this uniqueness, we need to take a look at both distinctions which may have motivated Hume to draw the natural/artificial distinction in the *Treatise*. After this, I will compare Hume's treatment of this distinction in the *Treatise* and the second *Enquiry* and explain their differences by changes in Hume's concerns and philosophical methods in this later work.

Let me start with Pufendorf's distinction between 'natural', or 'connate', and 'adventitious' obligations that Buckle highlights. The ground of this distinction lies in their causal origins:

The former [connate obligations] are those that inhere in all men by virtue of the fact that they are animals endowed with reason, or that accompany rational nature itself, as such ... *Adventitious obligations* are those placed upon men, with their express or presumed consent, because of an antecedent human deed. (DJNG 3.4.3; 167)

To recall, obligations, for Pufendorf, are 'moral qualities' of obliged persons which are imposed by the wills of intelligent beings. Their 'authors', so to speak, can be both God and human: we lie under connate obligations simply as free rational creatures subjected to God's legislation, but as such creatures, we can also create adventitious obligations by our acts of will. This distinction accords with one out of three senses of '*the natural state of men*' in Pufendorf's theory: connate obligations are those that human beings have in a state where they relate to each other merely "on the basis of the simple common kinship which results from similarity of

⁶⁰ Also see Haakonssen (1996, 118).

2.3 ‘Natural’ and ‘Artificial’ Obligations

nature”, a state “antecedent to any agreement or human action by which particular obligations of one to another have arisen” (*De Officio*, 2.1.5; 116). Representative examples of connate obligations are those of no injury as well as those to humanity and gratitude, and examples of adventitious obligations are those to respect others’ property and to obey governments.⁶¹ Importantly, for Pufendorf, adventitious obligations are also ‘natural’ insofar as they are also sanctioned by the natural law: human beings can create such obligations by pacts or promises because the natural law imposes on them a *general* and *indefinite* obligation to keep promises. By making a promise, one transforms this general obligation into a *particular* obligation to perform the *promised* action with one’s own moral power or authority (*potestas*) that God bestows on human beings.⁶²

The apparent similarity between Pufendorf’s and Hume’s distinctions lies in Hume’s view that what is ‘artificial’ depends on antecedent human actions and that the obligations to respect others’ property and obey civil governments fall in this category. Hume, in a footnote attached to the third Appendix of the second *Enquiry*, claims that his account of the origin of property is “in the main, the same with that hinted at and adopted by Grotius” and cites Grotius’s account at length (EPM App. 3.8n; 307). As Buckle observes, Grotius, Pufendorf, and Hume all hold that the rights and obligations of property were established by human actions gradually in history, and unlike Hobbes, all of them hold that this can be achieved without either instituting or subjecting to a sovereign. Pufendorf writes that property was established by agreements or *conventiones*,⁶³ and Hume explains this process by human conventions and uses ‘agreement’ as a synonym for ‘convention’ (T 3.2.2.10; 499). Also, for all authors, the invention of property was motivated by human beings’ natural need for social cooperation for their self-preservation and well-being and was enabled by their ability to lead a sociable life and inventing new social practices. As society is necessary for human beings, and property rules are necessary for the maintenance of societies, such an invention is not arbitrary and can be commonly observed across all human communities (T 3.2.1.19, 3.2.6.1; 484, 526).⁶⁴

What makes their theories significantly different, at the same time, is that for Hume, the convention or agreement through which property was first invented does not involve a promise

⁶¹ See DJNG 3.1.1, 4.4.4-6, 7.2.7-13; 158-59, 177-79, 211-14.

⁶² See DJNG 2.3.24, 3.4.1, 4.4.14; 156-57, 165-66, 185-86.

⁶³ See Pufendorf (1744, 518). Kennett sometimes translated *conventiones* as ‘compact’ and sometimes as ‘agreement’. See DJNG (1729) 4.4.4; 321.

⁶⁴ Especially see the discussion of Buckle (1991, chaps 1, 2, 5). For Grotius’s account of the origin of property, see DJBP 2.2.4-5; 426-27; for Pufendorf’s, see DJNG 4.4; 175-86.

2.3 'Natural' and 'Artificial' Obligations

(T 3.2.2.10; 499), and the 'natural law' about performing promises, by itself, was established by conventions and is, in this sense, 'artificial' (T 3.2.5.1; 516). Conventions do not work, therefore, by the force of a more general natural law but make the 'natural laws' forceful. For Grotius and Pufendorf, the state of nature in which human beings are supposed to invent property with pacts or agreements is a state in which they live under God's prescription and with rights and obligations in accordance with the natural law. For Hume, though property also emerged from a 'state of nature'⁶⁵, the state of nature is not a state with 'natural' rights and obligations: the ideas of right and obligation, to repeat, are dependent on conventions. On this point, Pauline Westerman (1994, 92–99) observes that though Grotius and Pufendorf, on the one hand, and Hume, on the other, all provide theories of the historical development of property, the former's accounts have two aspects: a *descriptive* aspect concerning how property was invented as a product of human socialization, and a *justificatory* aspect concerning how this process conforms with the natural law. Hume separates these two aspects and leaves the latter to his theory of moral sentiments. I think this is basically right, though, as I will argue later in Chapters 5 and 6, I regard Hume's account of the historical origin of justice as also an account of the history of our moral sentiments about justice. As Dario Castiglione (1987, 169) argues, for Hume, "society is not the product of normative processes, but, on the contrary, these necessarily follow from its establishment"; our moral judgements made in terms of right and obligation are products of some 'human artifice' or social coordination.

Another noteworthy distinction in natural law theories which has apparent similarities with Hume's distinction between natural and artificial obligations is the one between imperfect and perfect duties, a distinction held by Grotius, Pufendorf, and Hutcheson. More strictly speaking, for all three authors, this distinction rests on the distinction between perfect and imperfect rights by which the performance of such duties is claimed. Perfect rights are those that one holds with respect to one's life, limbs, liberty, property, and the performance of promises, while a representative case for imperfect rights is that which can be claimed based on humanity. For all three authors, the most significant difference between these rights is that perfect rights can be claimed by force, while imperfect rights cannot. These authors' grounds for this distinction are not exactly the same. For Grotius, perfect and imperfect rights are different types of moral qualities of a person which enable one to "have, or do, something justly". Perfect rights are 'faculties', while imperfect ones are called 'aptitudes'; the former signifies "which a Man has to his *own*", while the latter is one's merit or worth to be treated in a certain manner by those

⁶⁵ I will discuss Hume's conception of the state of nature in Chapters 4 and 5.

2.3 'Natural' and 'Artificial' Obligations

who have virtues like mercy or prudence.⁶⁶ Only perfect rights and the type of justice which responds to such rights are 'right' and 'justice' in their proper senses, according to which right, as a quality of actions, "consists in leaving others in quiet Possession of what is already their own, or in doing for them what in Strictness they may demand" (DJBP Preliminary. 10; 88-89). Pufendorf puts perfect and imperfect rights on a more equal footing and does not claim that they have different natures as moral qualities, though he also highlights perfect rights' distinctive relation to *suum* or what is 'one's own'. The reason why only perfect rights are enforceable, Pufendorf argues, is that the natural law's precepts about these rights "contribute to the being of society", while those about imperfect rights only contribute to the society's well-being. Also, because imperfect rights bind one "from some moral virtue" and through one's "sense of decency and conscience", and such virtuous motives cannot be extorted by external force, enforcement is pointless in such cases.⁶⁷ Hutcheson's ground of distinction is more similar to Pufendorf's: he calls both types of rights 'faculties' and distinguishes them by their respective necessity for promoting the public good: the universal violation of perfect rights "would make human Life intolerable", while imperfect rights only "tend to the improvement and increase of positive Good" in a society (*Inquiry*, 2.7.6; 183-84).

The apparent similarities between these authors' perfect/imperfect distinction and Hume's artificial/natural distinction consist in (i) an overlap of representative items in their catalogues, as rights and obligations about property and promises are 'artificial' for Hume, and benevolence is the 'natural' duty that he most frequently talks about, and (ii) similar distinguishing features of 'perfect' and 'artificial' obligations, as for Hume, the natural laws of property and promises can be *enforced* by civil laws, and the rights and obligations 'determined' by such laws share several characteristics of perfect rights in the Grotian tradition, like their *strictness*, *preciseness*, and their *negative* requirements – as Grotius argues, right in its proper sense only requires one to leave others in peaceful possession of what is their own, which does not need to involve any positive action.⁶⁸ The apparent dissimilarity between these distinctions, meanwhile, consists in their treatment of the obligations not to harm others' lives, limbs, and liberty; in other words, of *perfect* and *connate* obligations. Hume does not discuss these duties in his theory of justice, which is restricted to property and promises, and he takes 'property' to concern only external goods which are distinguished from "the internal satisfaction of our mind" and "the external

⁶⁶ DJBP 1.4.4-8; 138-47.

⁶⁷ DJNG 1.1.20, 1.7.7; 107, 131.

⁶⁸ For these similarities, especially see Schneewind (1998, 365–66) and Harris (2010a, 26–28, 35–36).

2.3 ‘Natural’ and ‘Artificial’ Obligations

advantages of our body” (T 3.2.2.7; 487). Neither does he say that we have ‘rights’ over our lives and limbs; in the limited texts where he talks about rights, this concept is either juxtaposed with ‘property’ and ‘obligation’, claiming that all these three ideas depend on conventions, or discussed in contexts about particular rules of property or allegiance.⁶⁹ As Harris (2019, 218–20) observes, the lack of a place for connate obligations in Hume’s theory of justice also makes up one important difference between Hume’s conception of justice and the meanings of justice adopted by natural law authors or Hume’s contemporaries under the influence of this tradition.

There are two strategies to find a place for such rights and obligations in Hume’s taxonomy. One is to suggest that the obligation not to harm others’ lives and limbs is ‘natural’ because it is subordinated to the obligation of benevolence, especially based on the natural mechanism of sympathy and the passions of love and pity, which, when working properly, may prevent people from inhumane treatment of others. The existence of such natural inclinations may explain why, in Hume’s view, the principal difficulties for people in a state before the invention of justice to live a social life consisted in the lack of stability of external possessions rather than, like in a Hobbesian state of nature, the fear of violent death under others’ hands (T 3.2.2.7, 3.2.2.12; 487–88, 491). The other strategy, as suggested by Rosalind Hursthouse (2009, 269–72), is that connate perfect rights are, like adventitious perfect rights, ‘artificial’, and what can be said about property rights can also be said about one’s rights over one’s own life and body. Hursthouse’s worry about the first reading is that the principle of humanity cannot establish a domain of the ‘*suum*’ as the natural lawyers understand it; for instance, if one is convinced that killing another person will make them better off, one’s benevolence might encourage one to kill despite the victim’s disagreement. Also, when Hume talks about “the laws of humanity” about treating gently those “creatures” who are “rational” but “possessed of such inferior strength ... that they were incapable of all resistance”, he argues that such “laws” do not give such creatures any rights (EPM 3.18; 190), which seems to support Hursthouse’s view.

I do not think that this question can have a determinate answer with robust textual support. The difficulty with connate perfect rights and obligations arises from the appearance of Hume’s taxonomy: his natural/artificial distinction is similar to the connate/adventitious distinction in

⁶⁹ Hume never defines what a right (understood in the sense of subjective right or *ius*) is and what types of rights there are. He mentions right in this sense 60 times in the *Treatise*, all contained in 5 sections in Book 3, part 2, and almost all of them are discussed in either of these two manners (except for people’s right of resistance in T 3.2.10.16; 564, but he does not explain what is the nature and ground of this right). The case of the second *Enquiry* is similar, with one exception in a footnote where Hume remarks that one has a right to dispose of one’s life (EPM 7.9n; 253). Because Hume mentions this right as a premise for an “ancient maxim” that it is a duty not to prolong one’s life when it becomes dishonourable, it is unclear whether Hume himself endorses this point or not.

2.3 'Natural' and 'Artificial' Obligations

some respects but the imperfect/perfect distinction in other ones. Hume's footnote on Grotius suggests that it is hardly possible that he was unacquainted with these traditional distinctions, and a more beneficial question to ask here is *why* he invented this new distinction when he had more widely accepted categories at hand. It would be helpful to notice in the first place that, in my view, Hume's accounts of obligation were not composed out of a spirit of system: he might not have been interested in constructing a systematic theory of rights and obligations as Grotius, Pufendorf, or Hutcheson were in their jurisprudential works or moral philosophy textbooks, and Hume might not have been that much concerned with finding a neat match between his innovative thoughts and the established traditions. Hume's accounts were more question-driven, focusing on what was most salient when certain observations were put together and questions arose from them, and attempting to answer these questions with resources in his study of human nature. This is especially likely to be true for the *Treatise*, which may explain some structural features of this book which may appear strange from an architectonic standpoint since Hume, more than once, starts with more complex subject matters and proceeds to simpler ones (like putting indirect passions before direct passions, artificial virtues before natural virtues), appears 'careless' about definitions and taxonomies (like, unfortunately for this thesis, in the case of obligation), and pays significantly unbalanced attention to different categories in a distinction, or different items in a category, like in the cases of the obligation to justice and the obligation to have some natural virtues. Considered from this perspective, a helpful interpretative strategy is to first track Hume's questions and then his answers rather than presuming that he intended to construct and follow a complete conceptual map when writing on obligation.⁷⁰

The similarities between Hume's natural/artificial distinction and the connate/adventitious distinction on the one hand, and the imperfect/perfect distinction on the other, considered from this perspective, shed light on Hume's distinction by suggesting a probable source of questions and inspiration for answers. (i) The distinction between perfect and imperfect obligations raises a question for Hume because, as I discussed in Section 1.3, the basic landscape of our moral experience, as Hume finds it, is a world of approbation and disapprobation, virtues and vices, but some types of moral judgements, namely, those about *perfect* rights and obligations, have characteristics which diverge from the manner in which we judge about virtues and vices. (ii)

⁷⁰ A similar interpretative strategy is adopted by Harris (2019, 233), who argues that Hume's peculiar definition of justice can be explained by his critique of Hutcheson's theory of moral sense as a special mental faculty and the consequent need to explain the origin of justice with a conjectural history; I largely agree with this interpretation. Forbes (1975, 86–89) also maintains that Hume neglected connate perfect rights in his theory of justice because these topics were not his dominant concern, though he reads this 'dominant concern' as being the *economic* aspect of society.

2.3 'Natural' and 'Artificial' Obligations

This question is more salient regarding *adventitious* perfect rights and obligations since while the obligation not to harm others' lives and limbs can, at least, find *some* ground in human beings' natural passions and evaluative mechanism, the obligations to respect others' property and perform promises cannot. (iii) Finally, Hume's answer to this question was likely inspired by the narratives of the state of nature and human agreements that previous authors developed to account for the distinction between connate and adventitious obligations, though he gave this narrative a naturalistic and sentimentalist turn which I will discuss in Chapter 4.

Interestingly, these three aspects in which Hume's distinction responds to previous theories also correspond to his three different characterisations of the differences between natural and artificial virtues in the *Treatise*. (i) The differences between perfect and imperfect obligations are highlighted in two arguments in *Treatise* 3.2.6 where Hume discusses the preciseness and strictness of our moral judgements about the rights and obligations of justice (T 3.2.6.7-10; 529-33). (ii) What is more distinctively problematic about adventitious perfect obligations, that is, their lack of ground in natural passions, is addressed in Hume's argument for the artificiality of justice in *Treatise* 3.2.1 and another in *Treatise* 3.2.6 (T 3.2.6.2-6; 527-29), and his statement in *Treatise* 3.3.1 that "the only difference betwixt the natural virtues and justice lies in this, that the good, which results from the former, arises from every single act ... Whereas a single act of justice, consider'd in itself, may often be contrary to the public good" (T 3.3.1.12; 579). (iii) Finally, Hume's initial explanation of what he means by 'natural' and 'artificial', i.e., the latter depends on human designs and projects (T 3.1.2.9; 475), is similar to what previous authors meant by what is natural and adventitious, although they have different understandings of how human actions produce differences in the moral domain. These groups of characterisations of the differences between natural and artificial virtues, as will become clearer in the following chapters, are related to each other in the *Treatise* as *explananda* and *explanation*: the influence of human actions, through convention and education, explains the 'artificial' manner in which just actions are regarded as useful, and this manner works to explain the features of the rights and obligations of justice as perfect rights and obligations.

In this explanatory project, connate perfect obligations are left out, probably because they are not as problematic as adventitious perfect obligations considered against the background of Hume's virtue-centred moral theory. It cannot be inferred from the absence of such obligations in Hume's theory of justice that Hume does not take them as sharing some features of perfect obligations, like being strict and precise, no matter whether they are natural or artificial. More generally speaking, as the distinction between natural and artificial obligations is more likely

2.3 'Natural' and 'Artificial' Obligations

driven by Hume's specific concern, despite its apparent similarities with the imperfect/perfect distinction, it does not, or there is no determinate textual evidence to support that it does, trace other aspects of the perfect/imperfect distinction. Hume does not suggest that considerations of artificial obligations should always *precede* those of natural obligations when they conflict in particular situations.⁷¹ Hume does not make the point that the performance of artificial duties is more *necessary* for maintaining society than all other duties, as the duties of parental care and humanity, as argued by Shaver (1992, 552–55), may be as necessary as the duty of justice. One cannot tell, correspondingly, whether Hume would take all perfect obligations as 'artificial' based on their priority or necessity. At the same time, I think Hursthouse is right in speculating that Hume might agree that some aspects of connate perfect rights cannot be reduced to 'the laws of humanity'. Hume might agree that such rights are as 'artificial' as the right of property, and the same model of historical origin may be applicable to these aspects as well, though probably with alterations. Our moral disapprobation of actions which violate such rights, if there is a place for such rights in Hume's theory, may, in a psychological respect, contain a *mixture* of the disapprobation of a natural vice (like cruelty) and that of the violation of an artificial right (like the right to dispose of one's own life, as Hume mentions in pass in EPM 7.9n [253]). This part of Hume's theory, or what may be developed as a 'Humean' theory according to Hume's principles, however, can only be a subject of conjecture just like many other topics about right and obligation on which Hume did not give determinate words in his works.

Finally, it is noteworthy that the distinction between natural and artificial virtues ceased to be Hume's dominant concern in the second *Enquiry*. He does not mention this distinction in the main-body sections of this work, and he comments in its third Appendix that "it seems vain to dispute whether justice be natural or not" (EPM App.3.9; 307). This does not indicate, however, that he changed his substantial view concerning the difference between justice and other virtues in this later work. He does refrain from using the words 'natural' and 'artificial', and in the only paragraph about this distinction, he puts his stress on the naturalness of justice, i.e., that it arises from human beings' natural faculties and is commonly found in all societies as it is necessary for maintaining a social life (EPM App.3.9; 307). In the footnote attached to this paragraph, he clarifies that when 'natural' is opposed to 'artificial', where the latter means

⁷¹ Some scholars hold the opposite view. Jacqueline Taylor (1998, 21), for instance, suggests that for Hume, "the conventions of justice determine the scope of the natural virtues". Taylor also holds that Hume's distinction between artificial and natural virtues corresponds to that between perfect and imperfect duties, though her point of reference is Kant.

2.3 'Natural' and 'Artificial' Obligations

that something supposes “reason, forethought, design, and a social union and confederacy among men”, justice can be said as artificial (EPM App.3.9n; 307-308). This position does not differ from that in the *Treatise*, as I discussed above, though it is addressed in a much weaker tone and given a much lesser place.

What changed in the *Enquiry* concerning the natural/artificial distinction, however, went further than a mere change in the choice of terms and tones. Two other major differences can be observed between these works in this respect. First, regarding Hume’s observations of the phenomena about this difference, or the two aspects which he takes as his *explananda* in the *Treatise*, as I discussed above, he cuts down some parts of these observations but preserves the other parts in the *Enquiry*. He dismissed the argument on the motive for justice in *Treatise* 3.2.1 and that on the preciseness of the rights and obligations of justice in *Treatise* 3.2.6. What is preserved is the difference in the manners in which people evaluate the benefits of justice and other publicly useful virtues, namely, the good that results from the latter “is in itself complete and entire”, but the benefits of justice are “not the consequence of every individual single act; but arises from the whole scheme or system concurred in by the whole, or the greater part of the society” (EPM App.3.2-3; 304). Hume makes this difference the topic of the third Appendix, and in this way, gives it a *more* prominent place compared to the *Treatise*: though the same point is mentioned more than once in the *Treatise*, it is subordinated to other topics, that is, the conventional origin of justice and the origin of moral sentiments (T 3.2.2.21-22, 3.3.1.12; 496-98; 579) and may not attract his readers’ attention as much as other points on the difference between justice and the natural virtues. Two other arguments in the *Treatise* are also retained, though with alterations: the one on the strictness and universality of the rules of justice (T 3.2.6.9; 531-33) is mentioned in Appendix 3, as an explication of the difference between the evaluation of single acts and the whole system (EPM App.3.6; 305), and the argument on the definition of property (T 3.2.6.2-6; 527-29) is rendered into four paragraphs which Hume adds to Section 3 in the 1753 edition, arguing for the point that the rules of justice are derived not from an original instinct but from reflections on public utility (EPM 3.40-43; 202-202). In short, I think Hume maintained most of his observations on the so-called ‘artificial’ characteristics of justice in the *Enquiry*, though he put different emphases in these works, especially by dismissing the argument on the motive for justice and stressing the contrast between the ways in which justice and other virtues are regarded as useful. As I will argue in Chapters 5 and 6, I read the latter difference as central to Hume’s natural/artificial distinction in the *Treatise*, more central than the argument on motive. It is arguable, then, that Hume had a better sense of the

gist of his distinction in the *Enquiry*, but this shift of focus may also have been motivated by other changes between these works, as I will discuss later.

The other major change can be observed in Hume's accounts of the origin of justice. In the *Treatise*, he wrote a long section on the history of property and justice and followed it with two further accounts of the origin of the rules concerning the transference of property and promises. He also wrote a long section on the origin of particular property rules, especially addressing the psychological mechanisms which prompted such rules, like the associations of perceptions and the imagination's inclination to create new associations based on existent ones. These four sections connect two sections on the difference between natural and artificial virtues, and they together make up a complete unit of genealogical explanation (*Treatise* 3.2.1 to 3.2.6). In the *Enquiry*, these historical and psychological accounts are made in a much-abridged manner and dispersed to different parts of his sections. Hume offers *three* separate passages about the origin of property: one short passage concerning the gradual development of the rules of justice in Section 3 (EPM 3.21; 192); two short paragraphs on the point that justice arises from human conventions in Appendix 3 (EPM App.3.7-8; 306-307); and his quotation of Grotius's account in a footnote in Appendix 3 with a claim of agreement (EPM App.3.9n; 307). The psychological principles underlying particular property rules are only hinted about in passing in Section 3 (EPM 3.10; 195-96) and occupy a paragraph and a footnote in Appendix 3 (EPM App.3.10, 3.10n; 309-10). There is no separate account of the origin of the rules of transference and promises, and no remark on the order in which these practices arose. Most importantly, these discussions are made more like extended remarks that follow from Hume's main theses regarding justice: the gradual development of justice in Section 3 is used to support his thesis that "public utility is the *sole* origin of justice" (EPM 3.1; 183), and the conventional origin of justice in the third Appendix is discussed as what is *implied* by the difference between justice and other publicly useful virtues: *as* justice is useful on the scale of "the whole scheme or system" of actions on a societal scale, it should be acknowledged that justice arises from conventions or a "sense of common interest" (EPM App.3.7; 306). In other words, in the second *Enquiry*, Hume detached his central theses on the utility of justice and its special feature from his genealogical accounts; though these accounts can still be used to explain the features of our moral approval of justice, as may be evinced in Hume's comments on the conventional origin of justice, he no longer made his main theses depend on them.

When these differences are put together, what may be observed is a change in both Hume's primary concerns in his theory of justice and his methods to address his primary concerns. The

2.3 'Natural' and 'Artificial' Obligations

need to accommodate the rights and obligations of justice, with the features of adventitious and perfect rights and obligations, no longer occupied his central attention in the second *Enquiry*, and he no longer used genealogical explanation as a principal method to account for the moral merit of justice. This may make up a more substantial reason for Hume's downplay of the significance of the natural/artificial distinction in the *Enquiry*. This reason, nevertheless, does not imply that Hume no longer regarded the distinctive features of the obligation to justice as important, and whether this suggests that Hume became dissatisfied with his genealogical explanation in the *Treatise* because he found some difficulties in it by that time. In the first place, Hume may have had two *other* important reasons to marginalize his explanatory project regarding the obligation to justice in the second *Enquiry*: a change in the focus and design of this later work, and a change in his philosophical interests in general.⁷² Compared to Book 3 of the *Treatise*, the second *Enquiry* is organized with a more focused theme: to examine what qualities are approved as virtues or 'personal merits' and to explain their moral value by their usefulness or agreeableness; what is not so closely related to this central theme is moved to the appendixes, though this does not suggest that the latter topics are regarded as unimportant. In the case of justice, the emphasis of Section 3 is why justice is approved based on public utility, and almost all discussions on the difference between justice and other virtues are postponed to Appendix 3, probably to avoid distraction on the side of the readers. At the same time, Hume was less interested in giving psychological explanations of moral experiences in the second *Enquiry*, as may be evinced in his less technical account of sympathy there. As a considerable part of his genealogical explanation of the obligation to justice in the *Treatise* is made in terms of the association of perceptions, as I will discuss in later chapters, Hume's change of interest may have motivated him to pay less attention to this topic in general. Moreover, since the omission of his explanatory account of the obligation to justice in the *Enquiry* cannot be used as evidence for Hume's dissatisfaction with this account after publishing the *Treatise*, and since there is no explicit textual sign of disagreement between these two works in this respect, a more helpful approach to speculating about Hume's attitude about his earlier accounts is to consider these accounts' consistency with his points made in the *Enquiry*, and, in general, I do not find them inconsistent. Detailed considerations, of course, can only be made when I examine Hume's explanatory project in the *Treatise* later.

Because I take the need to accommodate the rights and obligations of justice in his virtue-centred moral theory as a central theme for Hume's theory of obligation, these changes between

⁷² For these differences, especially see the discussion of Harris (2015, 253–58; 2020).

2.4 Conclusion

the *Treatise* and the second *Enquiry* determine that I will take the former work as my primary point of reference in the following parts of this thesis, probably against Hume's own will as he declares in the Advertisement for the *Enquiries* that "the Author desires, that the following Pieces may alone be regarded as containing his philosophical sentiments and principles" (1975, 2). This does not mean that I take the second *Enquiry* (and other later works of Hume, like the *Essays*) as unimportant, containing no significant new thoughts, or that I presume *a priori* that all these works are in alignment with each other. On the contrary, there are many important discussions about duty, obligation, and justice in the second *Enquiry* which were not made, or not made in a manner as clear, precise, or comprehensive, in the *Treatise*. I will examine these discussions and compare them with the *Treatise* accounts in the following chapters, and where I find some remarks in the *Enquiry* consistent with the general position Hume makes on the corresponding topic in the *Treatise*, I will suggest that these texts may be read together to shed light on some aspects which he did not address very clearly in one of these works. Above all, meanwhile, the differences regarding Hume's primary concerns, philosophical methods, and his interest in the distinction between natural and artificial obligations in these two works should be kept in mind; these differences will prove to be important, for instance, when I come back to Hume's method of genealogical explanation in Chapter 5.

2.4 Conclusion

This chapter is an attempt to disentangle Hume's taxonomy of obligations, especially that made in the *Treatise* in terms of what is 'natural' and 'non-natural', and to compare his distinctions with apparently similar ones maintained by previous authors, especially natural law theorists. I argued that Hume's distinction between 'natural' and 'civil' obligations corresponds to that between the obligations to observe natural and civil laws, though for Hume, natural obligation in this sense does not depend on any legislator. The distinction between 'natural' and 'moral' obligations, via Hutcheson's distinction which is the most probable direct source of influence for Hume, indirectly corresponds to the distinction between 'natural' and 'moral' aspects of the obligation to observe the natural law suggested by Cumberland and Pufendorf. These aspects are respectively the necessitating force of obligation, which works through human beings' desire for what is beneficial or naturally good and fear for what is naturally evil, and the moral values of obligatory actions which depend on God's legislation. For all these authors, including Hutcheson in his later works, the principle of unity of these aspects of obligation resides in the unity between God's attributes in his legislative acts. For Hume, who denies the natural laws' origin in a divine legislator, the natural obligation to observe the natural laws is a partial *cause*

2.4 Conclusion

for the corresponding moral obligation, and the sentiment of moral obligation can only become ‘binding’ or ‘necessitating’ with the help of further human artifice, namely, education. Finally, Hume’s distinction between ‘natural’ and ‘artificial’ obligations responds to both the distinction between ‘connate’ and ‘adventitious’ obligations and the one between ‘imperfect’ and ‘perfect’ obligations. Instead of being an accurate counterpart for these divisions, the natural/artificial distinction responds to them as Hume’s answer to questions which arise from them when they are to be integrated into Hume’s virtue-centred moral theory.

This comparative attempt also helps, in my view, to put different aspects of Hume’s conception of obligation into one complete line of thought, if not one structured framework. As I discussed at the beginning of Section 2.1, the distinctions between natural and non-natural obligations make up one of the most prominent features of Hume’s account of obligation in the *Treatise*, and these distinctions, to a certain extent, dominate all his arguments on this topic. I also suggested that one probable reason for this is that Hume’s interest in the topic of obligation is intertwined with his concern with the relationship between human nature and morality. This relation can be most prominently observed in Hume’s distinction between natural and artificial virtues and obligations, and the natural/moral and natural/civil distinctions, in the *Treatise*, are discussed as a part of his explanatory project about artificial obligations: the former distinction is only applied to artificial obligation, and civil obligation is a subset of artificial obligation. Though these distinctions do not fit neatly under one conceptual framework, they can all be located in Hume’s causal explanation for the origin of social and political orders. When these distinctions are compared to the distinctions in natural law theories, one may especially observe a shift from a framework based on a legislator’s decree to another based on the natural development of social order and its corresponding form of morality, which, in my view, may most helpfully capture what Hume meant to achieve with his account of obligation. As we now have this larger picture in sight, let me offer more details about it, and the best place to start is with what Hume means more exactly by ‘moral obligation’.

Chapter 3

Unpacking the Sentiment of Moral Obligation

The two previous chapters were an attempt to delineate the basic features of Hume's conception of moral obligation and locate it in its intellectual context, especially in relation to natural law theories which, as I argued in Chapter 2, were an important interlocutor that Hume may have had in mind when developing his account of obligation in the *Treatise*. This account, which is interestingly structured with three distinctions between 'natural' and 'non-natural' obligations, presents a framework of the gradual development of social order and moral conceptions out of the natural faculties and inclinations of human beings, a process which finally led to 'our' moral experience in which various forms of obligations direct our practices and impose restraints on our natural motives. This framework, in Hume's moral theory, substitutes the more traditional framework of divine legislation in which obligation is defined and divided into different types or aspects.

From this chapter on, I dive into more details concerning Hume's understanding of moral obligation and his causal account which explains its features and its place in his virtue-centred moral theory. The best place to start is his delineation of moral obligation in *Treatise* 3.2.5.4. This is not a full-fledged definition as it lacks clarification in some important aspects, which makes it insufficient for distinguishing moral obligation from other concepts, and I will, by resorting to other texts, try to work out a finer-grained picture of Hume's conception of moral obligation. I will analyse Hume's rudimentary delineation of moral obligation and supply this text with further considerations in four aspects, namely, regarding (i) its features as a perception, (ii) its causal origin in a psychological aspect, (iii) its typical objects, or the species of the things regarded as morally obligatory, and (iv) its scope, especially when compared to the approbation of virtues. Section 3.1 focuses on the first aspect, where I argue that the sentiment of moral obligation differs from the approbation of virtue for its constitutive emotion, which is a painful feeling, and it differs from other forms of moral disapprobation because it is typically caused by ideas of the absence of a mental quality or the non-performance of an action. In Section 3.2, I argue that the sentiment of moral obligation arises either when the lack of a virtue is regarded as vicious, or when the disapprobation of a vice is associated with the idea of some remedy for

3.1 Obligation, Virtue, and Vice: A Preliminary Analysis

this vice. Moreover, I consider Hume's claim that '[o]ur sense of duty always follows from the common and natural course of our passions' (T 3.2.1.18; 484) and discuss three explanations for this claim, i.e., (i) deviation from the common course of human nature is, by itself, vicious; (ii) the sense of duty is always formed by comparison with the common course of human nature because it is a self-centred indirect passion; and (iii) the sense of duty is guided by the common course of human nature because it is guided by people's general expectations for an averagely virtuous character. I argue against the first two explanations, and I suggest that the sensation of disappointment upon one's failure to live up to some general expectations may be a part of the psychological origin of the sentiment of moral obligation. Finally, I examine the typical objects of moral obligation in Section 3.3. Because the sentiment of moral obligation arises, in many cases, as a disapprobation of one's lack of a virtuous quality, the objects of moral obligation, or the contents of one's duty, in such cases, are the virtuous qualities that one is expected to have. While some such qualities are constituted or signified by virtuous *motivating passions* such as benevolence and gratitude, I argue, against a popular reading, that some other morally obligatory qualities can be characterised as one's *abilities* or *habits* to act in a specific manner rather than motives. Based on Hume's theory of moral sentiment, it is questionable whether merely external actions, specified without referring to an agent's motivational status, can be proper objects of moral obligation; this question will become an important concern when we turn to the moral obligation to justice. Regarding the scope of moral obligation, there is no firm textual ground for distinguishing it from the scope of virtues. I argue that the principle 'ought implies can' cannot provide a standard for judgements about moral obligation since Hume does not endorse this principle on the level of individual agents, but it is plausible that judgements about moral obligation are more strongly influenced by considerations about one's social roles.

3.1 Obligation, Virtue, and Vice: A Preliminary Analysis

As I noted in Chapter 1, a crucial question which confronts Hume's interpreters concerning his conception of moral obligation, understood as a type of sentimental response to character traits and actions, is how it relates to other moral sentiments: Darwall (1995, 296) holds that moral obligation is the same as the approbation of virtue; Shaver (1992, 547, 555) maintains that it is the same as the disapprobation of vice; and Baier (2010, 130–35) suggests that it differs from both of these sentiments. In this section, I argue against the first two views by examining Hume's rudimentary delineation of judgements about virtue and moral obligation in the *Treatise*. In my reading, the sentiment of moral obligation is a specific type of disapprobation

which differs from other types by the structure of the ideas which typically cause it and arise from it.

To start, let me cite again the only passage in Hume's works where he explains what moral obligation is in general:

All morality depends upon our sentiments; and when any action, or quality of the mind, pleases us *after a certain manner*, we say it is virtuous; and when the neglect, or non-performance of it, displeases us *after a like manner*, we say that we lie under an obligation to perform it. (T 3.2.5.4; 517)

The function of this passage, in its original context, is to readdress Hume's sentimentalist view concerning the basis of moral judgements, and this serves as a premise in his argument for the point that a promise cannot, without the help of conventions, create a new moral obligation as it cannot change people's moral sentiments about a specific action. This passage does not count as a fully-fledged definition of virtue or moral obligation. It is helpful to compare Hume's delineation of virtue here with his two definitions of virtue in the second *Enquiry*: virtue, or personal merit, "consists altogether in the possession of mental qualities, *useful or agreeable to the person himself or to others*" (EPM 9.1; 269); and virtue is also defined in terms of "*whatever mental action or quality gives to a spectator the pleasing sentiment of approbation*" (EPM App.1.10; 289). It is controversial how these two definitions relate to each other; on this point, I agree with the readings of Don Garrett (2014, 124–26) and Hsueh Qu (2024) that neither of these two definitions is more fundamental than the other, and that they are mutually complementary.⁷³ Moreover, I agree with Garrett's suggestion that these two definitions respectively specify two aspects of 'virtue', which is a "sense-based concept" as it is derived from the sentiment of moral approbation. The first definition specifies the "productive aspect" of virtue, i.e., what quality of things the sentiment of approbation "ultimately serves to detect", and the second definition specifies the "responsive aspect" of virtue, i.e., "the characteristic mental response that results from the activation of the sense" (2014, 124). Compared to these two definitions, Hume's words on virtue in the passage cited above neither specify what the pleasant sentiment is nor specify what types of mental qualities this sentiment responds to. In a like manner, his delineation of moral obligation neither specifies what the painful *sentiment* is from which we derive the conception of obligation nor specifies what *things* typically give rise to this type of sentimental response. It cannot work, correspondingly, to distinguish moral

⁷³ For a helpful summary of different interpretations, especially see Qu (2024, 2–4).

3.1 Obligation, Virtue, and Vice: A Preliminary Analysis

obligation from other notions like reproach based on interested reasons, which means that it cannot serve as a proper definition.

Based on the parallel of virtue and moral obligation in this passage, it is possible to make the message it conveys about moral obligation clearer. The “*certain manner*” Hume talks about here is later explicated in *Treatise* 3.3.1 with his theory of sympathy and the ‘general points of view’ from which spectators reduce bias in their sympathetic feelings. Because the sentiment of moral obligation is felt “*after a like manner*”, this sentiment should be one of disapprobation which arises from the same mechanisms, and this coheres with Hume’s sketchy passage on the moral obligation to justice (T 3.2.2.24; 499-500). Unfortunately, although Hume promises to his readers to “give a full and satisfactory account” of moral obligation after examining the natural virtues (T 3.2.2.23; 498), he never returns to this topic and never gives any systematic account of what qualities of things the sentiment of moral obligation responds to. He does not, in other words, give any organized catalogue of *duties*. Hume himself does not draw any clear-cut distinction between ‘duty’ and ‘moral obligation’. For the convenience of discussion, I use the word ‘duty’ for the objects of the sentiment of moral obligation or the *things* that one lies under an obligation to have or perform, and ‘obligation’ for this *sentiment* and the idea of this sentiment. One may derive from the passage above a ‘responsive definition’ of duty: it is ‘*whatever action or quality of the mind the neglect, or non-performance of which gives to a spectator the displeasing sentiment of disapprobation, or a sentiment of moral obligation*’. One cannot, however, directly derive from the same passage a ‘productive definition’ of duty like the four-branched definition of virtue. Based on Hume’s sentimentalist position, such a productive definition can be worked out by examining the causal origin of the sentiment of moral obligation because, ultimately speaking, external objects are regarded as ‘virtues’ or ‘duties’ only because the view of them “causes a pleasure or uneasiness of a particular kind” (T 3.1.2.3; 471) in the minds of human spectators; the moral properties of ‘virtuousness’ or ‘obligatoriness’ do not reside in character traits or actions themselves but are attributed to them based on people’s sentimental response to them felt from some ‘general points of view’.⁷⁴ As

⁷⁴ I bypass many controversies on Hume’s metaethics here. Basically, I agree with the non-realist and cognitivist interpretation defended by Cohon (2008b, chap. 4). This means that for Hume, moral properties like virtuousness and obligatoriness have no existence independently of human beings’ psychological reactions, and moral assessments are capable of truth and falsehood. For different readings, see, for instance, David Norton’s (1982, 109–13) realist interpretation according to which moral qualities like ‘virtue’ are objective correlates of moral sentiments, and Chamberlain’s (2020) emotivist interpretation according to which moral judgements express sentiments rather than beliefs derived from sentiments.

3.1 Obligation, Virtue, and Vice: A Preliminary Analysis

Hume does not provide a full account of the origin of moral obligation, there is no firm textual basis on which a ‘productive definition’ of duty can be drawn accurately.

One may suggest that one reason that Hume does not discuss the origin of moral obligation in detail is that he makes no distinction between virtue and duty, and he does, in some texts (like his argument for the artificiality of justice in *Treatise* 3.2.1), switch between the languages of virtue and duty freely as if they were interchangeable. In my view, however, this suggestion cannot be granted since first, Hume’s rudimentary delineation of judgements about virtue and duty does indicate that their *sentimental bases*, i.e., the approbation of virtue and the sentiment of moral obligation, have different apparent features. Second, based on such different features, the *psychological processes* which produce these sentiments may be different. Third, Hume’s flexible use of the languages of virtue and duty in some arguments is not a determinate proof that he identifies virtue and duty because, as I will argue in the next chapter, it is probable that he argues in this manner because the same premises can be applied to both virtues and duties. In other places, like another argument for the artificiality of justice in *Treatise* 3.2.6 (T 3.2.6.7; 529-30), Hume does distinguish between virtue and obligation. I will examine these arguments later in Chapter 4; in this chapter, let me focus on my first two differences.

The point that the sentimental bases of our evaluations of virtues and duties are different can be elucidated by comparing Hume’s delineation of the sentiments of approbation of virtue and moral obligation in the passage cited above. Both of them are moral sentiments; and moral sentiments, for Hume, are impressions of reflection which arise from antecedent impressions or ideas (T 2.2.1.1-3; 275-76). Taken by themselves, all such impressions are *simple* and cannot be analysed into parts, so are the ideas directly copied from them. One cannot, thus, define the sentiment of moral obligation or its corresponding idea by their composition. Instead, one can distinguish one type of reflective impressions from another type by their phenomenological characteristics and typical causes and effects (T 2.1.2.1, 3.1.2.4-5; 277, 471-73). Hume does give such information in his rudimentary delineation: for moral approbation, the antecedent ideas which typically give rise to it are ideas of actions or qualities of the mind, and the “sensation” or “peculiar emotion” which constitutes the “being and essence” of this sentiment (T 2.1.5.4; 286) is a specific type of *pleasure*. On the other hand, the antecedent ideas of the sentiment of moral obligation are the “*neglect, or non-performance*” of some actions or mental qualities, and its constitutive sensation is a type of *displeasure or pain*. Finally, though Hume does not mention it here, the antecedent ideas of the disapprobation of vice are also ideas of actions or mental qualities, and its constitutive sensation is displeasure (T 3.3.1.30; 591). The

3.1 Obligation, Virtue, and Vice: A Preliminary Analysis

sentiment of moral obligation, consequently, differs from the approbation of virtue by both its antecedent ideas and constitutive sensation.⁷⁵ It shares with the disapprobation of vice the same painful constitutive sensation. At the same time, what is unique about the sentiment of moral obligation is that its antecedent ideas involve a relation of negation or contrariety: the contents of such ideas are specified as the *absence* or *omission* of some mental qualities or actions. Not all disapprobations are caused by ideas of such a structure: vicious qualities can be specified either positively (like cruelty, see T 3.3.3.8; 605) or negatively (like irresolution, see T 3.3.1.24; 588). The sentiment of moral obligation, therefore, may be distinguished from other forms of painful moral sentiments by the characteristic of its antecedent ideas.

One further difference between the sentiment of moral obligation and the disapprobation of vice can be observed by considering the ideas that arise from such impressions. I agree with the reading proposed by Garrett (1997, 196–99) and Cohon (2008b, 105–6) that, after feeling a moral sentiment of approbation or disapprobation, one can ‘copy’ this impression into a less vivid perception, i.e., an idea of this sentiment, and this simple idea (as it is copied from a simple impression) can be combined with other ideas to compose a *complex idea* in which a specific mental quality is represented *as* virtuous or vicious. The complex idea of humanity as a virtue, for instance, is composed of the idea of the character trait of humanity (which can be characterised as one’s inclination to be motivated to help others when they are in need), which is the antecedent idea of the sentiment of moral approbation, and the idea copied from this latter sentiment. When this complex idea has a sufficient degree of force and vivacity, it counts as a moral *belief*, and this belief is the content of one’s moral *judgement* about this quality.⁷⁶ The complex idea which makes up the moral judgement may be called a ‘consequent idea’ of the moral sentiment. Now, when one disapproves of a vicious mental quality like ingratitude, the consequent idea which arises from this sentiment is ‘ingratitude is vicious’ or ‘someone is vicious for being ungrateful’. Meanwhile, when one feels a sentiment of moral obligation toward the absence of a mental quality or the omission of an action, “*we say that we lie under an obligation to perform it*” (T 3.2.5.4; 517, my italics). The consequent idea, expressed here by ‘we say’, is the idea that ‘someone lies under an obligation’ to the thing which is *negated* in

⁷⁵ To recall, Hutcheson argues that “[w]hen we say one is oblig’d to an Action, we either mean ... or, 2. *That every Spectator, or he himself upon Reflection, must approve of his Action, and disapprove of his omitting it, fi he considers fully all its Circumstances*” (*Essay and Illustrations*, 146). The idea of moral obligation, for Hutcheson, is derived from both the approval and disapproval of the moral sense, and Hume’s delineation of the sentiment of moral obligation differs from it. For Hutcheson, the twofold origin of moral obligation leads to problems regarding the strength of moral obligation and how moral obligation correlates to rights; detailed discussions about these problems exceed the scope of this thesis.

⁷⁶ For a different view, however, see Chamberlain (2020, 1063–72).

3.2 Nature and Expectation: The Origin of Moral Obligation

the antecedent idea of the sentiment: the absent mental quality or the neglected action, and it is this negated part of the antecedent idea that the moral property of ‘obligatoriness’ is attributed to. This feature of its typical causes and effects, in my view, suffices to distinguish the sentiment of moral obligation from other forms of moral disapprobation on an apparent level.

Based on these apparent differences, one can either take the sentiment of moral obligation as a moral sentiment which is *sui generis*, different from both the approbation of virtue and the disapprobation of vice, or take it as a *subset* of the disapprobation of vice which has a specific causal context because it arises from a specific type of psychological process. Hume does not rule out the first possibility, but he would be less likely to embrace it because he observes the economic principle in his investigation of the origin of the moral notions: he aims at finding out causal principles which are as simple and general as possible (T 3.1.2.6; 473). Also, there is no evidence that the sentiment of moral obligation, apart from the structure of its antecedent and consequent ideas, has any special feature which cannot be observed in other forms of moral disapprobation. As I noted in Chapter 1, Hume does not share Kames’s view that the sense of duty is different from both moral approval and disapproval because it has an intrinsic authority, a “consciousness of necessity” which other types of moral sentiments lack (*Principles*, 31); for Hume, the stronger motivational efficacy of the sense of duty is fabricated by education. It is more plausible, then, that the sentiment of moral obligation is a subset of disapprobation which differs from other forms of disapprobation due to its special causal origin.

How, then, does this specific form of moral disapprobation arise? Though Hume does not develop a full account of this process, some of his accounts of duty and obligation, especially his remarks on natural (non-artificial) duties in his arguments for the artificiality of justice and his sketchy account of the origin of the moral obligation to justice, indicate that this sentiment is mostly plausibly derived from a previous approbation of virtue or disapprobation of vice. This will be the subject of the next section.

3.2 Nature and Expectation: The Origin of Moral Obligation

I argued in the last section that for Hume, the sentiment of moral obligation is a specific type of moral disapprobation which can be distinguished from other types of disapprobation by the special structure of the ideas which cause it and are derived from it. In this section, I investigate how this specific type of disapprobation may arise according to Hume’s psychology. I suggest that two probable routes of generation can account for the structure of the antecedent ideas of the sentiment of moral obligation, namely, this sentiment arises either when one’s lack of some

3.2 Nature and Expectation: The Origin of Moral Obligation

virtue makes one's character morally deficient, or vicious, or when the disapprobation of vice is associated with the idea of a remedy for this vice. The question which arises here is whether any further conditions apply to make one morally obliged to possess a virtue or avoid a vice. At this point, some scholars draw on Hume's remark on the duty of gratitude to argue that one will feel a sentiment of moral obligation when one finds oneself short of a virtue which can be commonly found in human nature. Despite the texts in which Hume builds a close connection between duty and the common course of human nature, I argue that the latter consideration cannot directly make the absence of a virtue vicious. Rather, it influences people's judgements about moral obligation by providing a guideline for general expectations for virtues, and one's failure to meet this expectation partly contributes to the generation of the sense of duty.

As the sentiment of moral obligation is a type of moral disapprobation which arises "*after the same manner*" (T 3.2.5.4; 517) as the approbation of virtue, I agree with Shaver (1992, 547) that a helpful starting point for investigating its causal origin is to consider the disapprobation of vice which arises when one considers a character trait which is either harmful or immediately disagreeable to either the person with this trait or other people, a sentiment which arises, like the approbation of virtue, through one's sympathy with the actual or potential victims of such a trait when one surveys its harmful or disagreeable tendencies from a 'general point of view'. Moreover, as the sentiment of moral obligation differs from other types of moral disapprobation by the structure of its antecedent idea, which involves a relation of contrariety, i.e., the non-existence of a quality of an action, one strategy is to consider the vicious qualities which are characterised as the absence of another quality, such as Hume's example of "inconstancy and irresolution; or a want of address in the management of men and business" (T 3.3.1.24; 587-88) which are destructive to one's own interest. The qualities which are absent in such a person's character, all being qualities in Hume's catalogue of virtues useful to oneself (T 3.3.4.7; 610-11), are then probable objects of the sentiment of moral obligation based on Hume's delineation of this type of moral sentiment.

Apart from such cases, I suggest that another way in which a sentiment of disapprobation may conform with Hume's delineation of the sentiment of moral obligation is when it arises toward a vicious character trait and is associated with an idea of a *remedy* for this vice, namely, another trait or some action which can usually prevent one from acting viciously. The originally disapproved vice, in this case, does not need to be characterised as the absence of a specific quality initially. It could be, for instance, one's inclination to feel excessive angry passions which "rise up to cruelty" and "form the most detested of all vices" (T 3.3.3.8; 606). When an

3.2 Nature and Expectation: The Origin of Moral Obligation

evaluator, with the help of their experience about human nature and practices, associates this vice with qualities or actions which can usually prevent one from acting out of such a vicious character, their disapprobation of the vice can be transformed into the disapprobation of the vicious agent's *failure* to employ this remedy, and this remedy, accordingly, is considered as an object of the sentiment of moral obligation, or a duty, of the agent under concern.

These two probable routes of generation are so far only hypotheses based on the features of the sentiment of moral obligation which need to be tested with textual evidence. Also, it is not clear, merely based on these suggestions, whether the absence of all virtues is to be regarded as vicious, and whether the disapprobation of all vices brings with itself an idea of an obligation to avoid this vice. It remains to be examined whether any further conditions or restrictions apply for a disapprobation to count as or generate a sentiment of moral obligation.

Concerning this question, one passage which has attracted many commentators' attention is Hume's remark on the duty of gratitude:

When any virtuous motive or principle is common in human nature, a person, who feels his heart devoid of that principle, may hate himself upon that account, and may perform that action without the motive, from a certain sense of duty, in order to acquire by practice, that virtuous principle, or at least, to disguise to himself, as much as possible, his want of it. (T 3.2.1.8; 479)

Based on this passage, Haakonssen (1978, 7–9) and Cohon (2008b, 197) suggest that a sense of duty arises when one finds oneself lacking a virtue common in human nature. This accords with the first route of generation I suggested above, as the lack of a common virtue is, in this case, regarded as a vice, and this virtue is, on the same account, regarded as a duty. Apart from this passage, Hume also argues in several other places about the relationship between judgements about moral obligation and the common course of human nature. For instance, he argues that:

A man naturally loves his children better than his nephews, his nephews better than his cousins, his cousins better than strangers, where every thing else is equal. Hence arise our common measure of duty, in preferring the one to the other. Our sense of duty always follows the common and natural course of our passions. (T 3.2.1.18; 483-84)

This partiality, then, and unequal affection, must not only have an influence on our behaviour and conduct in society, but even on our ideas of vice and virtue; so as to make us regard any remarkable transgression of such a degree of partiality, either by too great an enlargement, or contraction of the affections, as vicious and immoral. (T 3.2.2.8; 488)

3.2 Nature and Expectation: The Origin of Moral Obligation

Also, in his essay “Of the Original Contract”, where Hume distinguishes between two types of moral duties, he argues that:

The *first* are those, to which men are impelled by a natural instinct or immediate propensity, which operates on them, independent of all ideas of obligation, and of all views, either to public or private utility. Of this nature are, love of children, gratitude to benefactors, pity to the unfortunate. (E-OC 33; 479)

These texts suggest that considerations about human beings’ natural instincts and inclinations have an important role to play in determining which virtues are considered as duties, but what exactly this role is?

One answer which can be first ruled out is that being an exception to the common course of human nature is, by this characteristic itself, *morally* deficient. Hume makes this point in his arguments against the moral rationalists. Someone may contend, he argues, that we can judge that gratitude is a duty based on the fact that “human nature is generally grateful”. Hume replies to this contention:

But human nature is not so generally grateful, as to justify such a conclusion. Or if it were, is an exception to a general rule in every case criminal, for no other reason than because it is an exception? (T 3.1.1.15n7; 462)

Some other reasons are needed to render an exception to the general condition of human nature vicious. Also, Hume argues that when the word ‘natural’ is used in the sense opposed to what is rare and unusual, “perhaps virtue will be found to be the most unnatural. At least it must be own’d, that heroic virtue, being as unusual, is as little natural as the most brutal barbarity” (T 3.1.2.10; 475). As many virtues are exceptions to the common course of human nature, it cannot be true that any deviation from this course is vicious. Neither does Hume embrace a teleological conception of nature according to which what is ‘natural’ for human beings also means what conforms with their designed ends and moral perfections.⁷⁷ What accords with or deviates from nature, then, does not have any direct implication about moral values.

Another possible answer to this question is that the common course of human nature works as a standard for comparison for the production of indirect passions like pride and humility. As Hume notes about the ungrateful man, this man ‘hates himself’ for his lack of grateful motives.

⁷⁷ In his letter to Hutcheson on September 17, 1739, Hume remarks that Hutcheson’s teleological conception of what is ‘natural’ “appears to me pretty uncertain & unphilosophical” and depends on solving a group of questions “which are endless, & quite wide of my Purpose” (L 13; 1:33).

Scholars like Brown (1988, 82) read this feeling of hatred as an indirect passion that one feels about one's own person. According to Hume's theory, strictly speaking, the proper name of this passion should be *humility* rather than self-hatred, since while humility is caused by a quality of oneself which can produce pain, hatred is caused by a quality of someone else: as Hume stresses, "[w]e may be mortify'd by our own faults and follies; but never feel any anger or hatred" (T 2.2.1.2; 330). The self-centred indirect passions of pride and humility depend heavily on one's comparison between oneself and others: these two passions only arise when "the agreeable or disagreeable object be not only closely related, but also peculiar to ourselves, or at least common to us with a few persons" (T 2.1.6.4; 291). If the sense of duty was constituted by a passion of humility, it would follow that only the lack of those virtues which can be commonly found could produce a sense of duty, because when a virtue is rare, one's lack of it would not be "peculiar to oneself", and humility would not arise on this account.

This suggestion, however, is also implausible. The sense of duty, or the sentiment of moral obligation, as I noted in the last section, is a sentiment of disapprobation which arises "*after a like manner*" as the approbation of virtue. Although Hume argues in several places that virtue and vice "must give rise to" one of the four indirect passions of pride and humility, love and hatred (T 3.1.2.5; 473, also see T 3.3.1.2-3, 3.3.5.1; 574-75, 614), I agree with the interpretation of Radcliffe (2018, 132) that the approbation of virtue and disapprobation of vice are not, by their own nature, types of indirect passions, but always give rise to indirect passions.⁷⁸ Virtue and vice always give rise to indirect passions because they are mental qualities closely related to ideas of persons, and because they can produce pleasant or painful feelings. This satisfies the conditions for the mechanism of "double relation" which gives rise to indirect passions (see T 2.1.5.1-5; 285-87), but moral sentiments, in this process, play the role of the "separate pain or pleasure" which triggers the mechanism of double relation (T 2.1.5.3; 285-86) and should be distinguished from the indirect passions as the products of this mechanism. On this account, if by the self-hatred of the ungrateful person, Hume means the sentiment of moral obligation, as Radcliffe (2018, 125-26) reads it, it cannot mean the indirect passion of humility. Otherwise, if the self-hatred means the indirect passion of humility, it should be read as a result of the sense of duty. One cannot, therefore, confound between the condition for the production of

⁷⁸ Páll Árdal (1966, chap. 6), Paul Russell (1995, chap. 6), and Ainslie (2007; 2015) read Humean moral sentiments as calmer forms of pride or humility, love or hatred, and Cohon (2008a) argues that Humean moral sentiments are indirect passions, but they make up a distinct category aside from these four types. I disagree with the former reading, but I think Cohon's interpretation is attractive and worth separate discussion beyond the scope of this thesis. For a helpful overview of this controversy, see the discussion of Åsa Carlson (2014).

humility and the condition for the production of the sense of duty; it is still questionable how the common course of human nature works in the latter process.

It could also be supposed that the common course of human nature simply *coincides* with the measures of duty in the cases of virtuous natural instincts like benevolence and kindness to children (T 2.3.3.8; 417)⁷⁹: it just happens that human nature is commonly endowed with natural instincts and inclinations the lack of which is especially harmful and disagreeable, and thereby strongly morally disapproved. Such a connection, however, cannot account for Hume's strong claim that "[o]ur sense of duty always follows the common and natural course of our passions" (T 3.2.1.18; 484) and the frequency in which he addresses similar points.

With all these options excluded, I suggest that one probable reading is that considerations about the common course of human nature guide the formation of the sense of duty by guiding people's general expectations about the character traits that an averagely virtuous person may have. Hume does not talk about people's expectations as a measure of moral obligation. In the second *Enquiry*, though, he remarks that expectation helps to explain why the absence of a virtue is regarded as vicious:

The absence of virtue may often be a vice; and that of the highest kind; as in the instance of ingratitude, as well as meanness. Where we expect a beauty, the disappointment gives an uneasy sensation, and produces *a real deformity*. (EPM 7.9n; 253-54, my italics)

His example here, like in the case of the man who hates himself, is also ingratitude. One's lack of the virtue of gratitude makes up a vice partly because of the feeling of disappointment which arises when one fails to meet an expectation for such a virtue; as disappointment is an uneasy sensation, it "produces a real deformity" in one's character *at least* by making one's character disagreeable to those who expect for this virtue. This disagreeableness may work together with the harmful tendency of one's lack of gratitude and make one's character more depraved when considered from a general point of view. Another example in which expectations influence moral evaluation may be found in the *Treatise*, where Hume remarks on "a certain indulgence due to human nature" with respect to the evaluation of disagreeable passions: though anger, in

⁷⁹ Gill (2006, 247–55), Ainslie (2007, 104–6), and Kate Abramson (2015, 350–51) argue that for Hume, all virtues are rare, and qualities simply specified by natural instincts, even being useful and/or agreeable, cannot be virtuous in a proper sense. I do not agree with this reading. One reason they appeal to is that for Hume, all virtues can cause pride and love, and only comparably rare qualities can produce pride (T 2.1.6.4; 291). Hume does not argue, however, that only rare qualities can produce *love* when we react to others' virtues. Also, Hume observes that though "the same object always produces love and pride," it seldom produces them "in the same proportion" (T 2.2.10.7; 392). Some qualities, like good nature and generosity, easily produce love in others but not pride in oneself (T 2.2.10.8; 392, also see Hume's distinction between lovable and awful virtues in T 3.3.4.2n; 608).

3.2 Nature and Expectation: The Origin of Moral Obligation

general, is disagreeable, not all angry passions are regarded as vicious due to their commonness, and those who feel anger in a degree lower than “the greatest part of mankind” are even praised as virtuous (T 3.3.3.7; 605). Also, Hume notes that “[w]e make allowance for a certain degree of selfishness in men; because we know it to be inseparable from human nature, and inherent in our frame and constitution. By this reflection, we correct those sentiments of blame, which so naturally arise upon any opposition” (T 3.3.1.17; 583). While the disappointment that arises from one’s failure to possess an expected virtue strengthens moral disapprobation by making one’s character disagreeable, the expectation of a vicious quality may, when it is met, probably through some agreeable sensation when one’s expectation, weaken the original disapprobation, though Hume does not dive into details on how the latter process works.

The influence of expectations provides a promising explanation for the role of the common course of human nature in the formation of the sense of duty, since it makes the conformity with or deviation from it morally relevant, but only indirectly. On this account, for instance, one’s transgression of the natural degree of partiality in confined benevolence (like loving some stranger more than one’s children) is regarded as a breach of duty not only because this quality might be, by itself, harmful, but also because confined benevolence with its natural degree of partiality is a useful and agreeable quality which is generally expected, and one’s failure to meet this expectation makes one’s character disagreeable when examined from a general point of view. Moreover, though in most of Hume’s remarks on natural or non-artificial duties, the sense of duty arises from the idea of the absence of a virtue, I think expectations based on the common course of human nature can also play an important role if one first disapproves of a vice (like cruelty) and reflects about remedies for this vice. As pity and benevolence, which arise from human beings’ natural inclinations, can, as “the common and natural course of our passions” works, prevent one from treating others cruelly, the expectation for this common remedy may make one feel a stronger disapprobation toward cruelty and relate to this disapprobation the idea that it is obligatory to act with a certain degree of benevolence even when one feels strong anger. The sentiment of moral obligation, in both cases, arises as a combined result of an original disapprobation and considerations of the common course of human nature; the latter both work to magnify the strength of moral disapprobation and guide the formation of the antecedent and consequent ideas of this sentiment by informing people what one can be expected to do in relevant circumstances.

If the common course of human nature influences the formation of the sentiment of moral obligation by guiding people’s expectations rather than by the naturalness of the qualities under

3.3 The Objects and Scope of Moral Obligation

concern, one important consequence is that the *same* mechanism can work for those useful or agreeable qualities which are acquired not by natural instincts but by custom and education, as long as certain social practices or institutions render such qualities sufficiently common in a community. When commenting on Hume's remarks on natural duties, Haakonssen argues that it is compatible with Hume's account that "the natural principles in the human mind have conventional expressions, and that these can vary from time to time, and from place to place" (1978, 9). This is plausible: natural passions like parental love and gratitude can be expressed in different, customary behavioural codes in different societies, and inclinations to express such passions in a specific manner can make up aspects of the expected virtuous character traits and be regarded as one's duties. This does not make such duties 'artificial' in the sense that justice is, as I will argue in later chapters: the differences between 'natural' and 'artificial' virtues and duties do not imply that 'natural' virtues and duties cannot be acquired by education and shaped by custom. For now, however, let me focus on the topic of moral obligation in general.

To sum up, Hume's account of the duty of gratitude and his remarks on the common course of human nature as a guidance for the sense of duty supports my suggestion that the sense of duty may arise when the lack of a virtue is regarded a vice, and a plausible explanation for the role played by considerations of human nature in this process is that they influence people's expectations about virtues. The other route of generation, which starts with the disapprobation of a vice and combines it with ideas of the remedy for this vice, especially when this remedy may be commonly expected based on human nature or custom, may suit better the origin of the moral obligation to justice, which I will discuss in more details in Chapter 6. Even if general expectations may be acknowledged as a part of the psychological process which gives rise to the sentiment of moral obligation, and the common course of human nature may thereby be regarded as a standard of duty, the question still holds with respect to whether any further conditions apply to the formation of the sentiment of moral obligation. I will consider some possible conditions in the next chapter.

3.3 The Objects and Scope of Moral Obligation

In the two previous sections in this chapter, I examined the nature and origin of the sentiment of moral obligation according to Hume's rudimentary delineation of this sentiment and remarks on the duty to have some virtues which are common in human nature. These discussions, still, do not suffice to provide a 'productive definition' of duty in Hume's theory because they cannot determine the *scope* of mental qualities and actions which are usually regarded as obligatory. I

3.3 The Objects and Scope of Moral Obligation

will approach this question in two aspects in this section. First, I discuss what types of things, in Hume's account, are the typical *objects* of moral obligation or are usually regarded as morally obligatory. Three optional answers may be given to this question, namely, (i) mental qualities or character traits, (ii) motives or motivating passions, and (iii) external actions which are specified independently of the motivational status of agents. Although some texts in the *Treatise* suggest that only motives are the proper objects of moral obligation, I argue against taking this restriction as universally applicable, as some duties may be qualities which are not characterised by any specific motives. At the same time, it is questionable whether external actions can be regarded as proper objects of moral obligation. Second, I consider Baier's two suggestions concerning Hume's view on the scope of duties, namely, that duties should be those qualities and actions which (i) can be voluntarily acquired or performed and (ii) are required based on the social roles or stations that one occupies. I argue that the latter suggestion is more plausibly in alignment with Hume's view compared to the former, but neither of them has firm support in Hume's text.

To start with the proper object of moral obligation, it may be recalled that Hume suggests that the idea of the neglect or non-performance of "any action, or quality of the mind" (T 3.2.5.4; 517) can give rise to a sentiment of moral obligation. Hume does not clarify, however, whether both *actions* and *qualities* of the mind are proper objects of moral obligation or proper items in a catalogue of duties, and whether they are regarded as morally obligatory in the same sense. Questions may arise here because Hume later claims that:

If any *action* be either virtuous or vicious, 'tis only as a sign of some quality or character. It must depend upon durable principles of the mind, which extend over the whole conduct, and enter into the personal character ... We are never to consider any single action in our enquiries concerning the origin of morals; but only the quality of character from which the action proceeded. (T 3.3.1.4-5; 575)

Since only durable mental qualities or character traits are proper objects of the approbation of virtue and disapprobation of vice, although an action can be evaluated as virtuous, it can only bear this moral property when taken as a *sign* or *expression* of a virtuous mental quality. An external action, not specified in relation to the mental quality which gives rise to it, therefore, cannot be a proper object of moral approbation or disapprobation. Hume's argument here only concerns the evaluation of virtues and vices, but if the sentiment of moral obligation, as I argued above, is derived from some previous approbation of virtue or disapprobation of vice, this restriction may also apply to evaluations of duties.

3.3 The Objects and Scope of Moral Obligation

In another piece of text in the *Treatise*, Hume suggests that the ‘ultimate objects’ of moral approbation and requirement are people’s motives:

’Tis evident, that when we praise any actions, we regard only the motives that produc’d them, and consider the actions as signs or indications of certain principles in the mind and temper ... When we *require* any action, or *blame* a person for *not* performing it, we always suppose, that one in that situation shou’d be influenc’d by the proper motive of that action, and we esteem it vicious in him to be regardless of it ... It appears, that all virtuous actions derive their merit only from virtuous motives, and are consider’d merely as signs of those motives. (T 3.2.1.2-4; 477-78, my italics)

Hume concludes this passage with a claim on virtues, but since he starts with observations on both moral approbation and requirement, what he infers from these observations should apply to both virtues and duties. His conclusion here, which I will refer to as the ‘Motive Principle’, is similar to his claim in *Treatise* 3.3.1.4 in suggesting that actions cannot be the proper locus of moral properties like virtuousness and obligatoriness. While *Treatise* 3.3.1.4 claims that the proper objects of moral evaluation are mental qualities, *Treatise* 3.2.1.4 claims that such objects are people’s motives. Hume illustrates the latter claim with the duty of parental care: “[w]e blame a father for neglecting his child. Why? because it shows a want of natural affection, which is the duty of every parent” (T 3.2.1.5; 478). Actions of parental care are only morally obligatory insofar as its typical motive, that is, the love of one’s children, is morally obligatory. The same can apply to the case of the ungrateful man who hates himself, which Hume mentions as another example after this passage: it is obligatory to perform grateful actions because it is obligatory to feel the “virtuous motive or principle” of gratitude (T 3.2.1.8; 479). Although one may think that one had fulfilled one’s duty in this respect by performing grateful actions, since such actions derive their moral value from the motive of gratitude, this person will still find his character deficient for his lack of this motive, and his self-disapprobation, or his sense of duty, will not be resolved. Merely performing grateful external actions, therefore, cannot make one actually fulfil one’s duty of gratitude; the same can be said about parental love.

Two questions arise here. The first is whether, despite Hume’s claims in these passages, it is still possible that certain external actions can be regarded as morally obligatory without being regarded as signs of an obligatory motive or quality. The second is how the first two types of objects of moral obligation relate to each other: whether it is consistent for Hume to claim that all actions derive their moral values from motives and that all actions derive their moral values from mental qualities. The first question primarily concerns the moral obligation to justice. As I will discuss in Chapter 4, one of Hume’s arguments on the distinction between justice and the

3.3 The Objects and Scope of Moral Obligation

natural virtues begins with the Motive Principle, where he argues that justice is not a natural virtue because the sense of duty about justice does not depend on any virtuous motive to justice. According to the interpretation of scholars like John Mackie (1980, 78–80) and Cohon (2008b, 171), to make this argument work, it is necessary to acknowledge that the moral obligation to justice takes external actions rather than motives or character traits as its objects. Also, Harris (2010a, 25–33) argues that Hume agrees with natural law theorists like Grotius that justice only raises requirements for external actions rather than motives. Whether these readings are plausible, and whether it is possible, given Hume’s theory of moral sentiments, for the moral obligation to justice to take merely external actions as its objects can only be answered in later chapters because a sufficient treatment of these questions requires examination of Hume’s argument on the artificiality of justice and his account of the origin of justice.

With the problematic case of justice set aside, another question about the objects of moral obligation, namely, whether they are *motives* or *mental qualities*, still persists. In my view, although Hume announces the Motive Principle (namely, “all virtuous actions derive their merit only from virtuous motives”) with a firm tone, and although many scholars read this principle as central to Hume’s moral theory, there are reasons to doubt whether this principle can be applied to all virtues, and the same reasons also generate doubts about whether this principle can be applied to all duties.⁸⁰ Concerning virtues, doubts may be maintained because not all virtuous traits of character in Hume’s catalogue are constituted or characterised by virtuous motives. To clarify, Hume uses ‘motive’ in the sense of motivating passions (see especially T 3.2.5.6, 3.2.6.9; 518, 532), which are token perceptions; character trait, on the other hand, is a quality of the mind which is “*durable* enough to affect our sentiments concerning the person” (T 3.3.1.4; 575). When Hume argues that actions are not proper objects of moral evaluation, his reason is that single actions, as transient events, do not *endure* long enough. Because of their short duration, pleasant or painful feelings which arise from ideas of single actions, when they are not considered in relation to a durable mental quality, cannot be constantly associated with ideas of persons (because such ideas are considerably more durable; see Hume’s argument in T 2.1.6.7; 293) and thereby cannot produce sentiments concerning persons, including moral sentiments. Motives, as particular perceptions or perception groups, are hardly more durable

⁸⁰ For scholars who take the Motive Principle as central to Hume’s moral theory, see, for instance, Darwall (1995, 288–96, 301–4), Don Garrett (2014, 264), and Geoffery Sayre-McCord (2016, 436–41). For doubts on whether Hume embraces this principle, see Baier (2010, 68). I argued that it should be doubted whether, in Hume’s view, this principle can apply to all virtues, or all natural virtues, in Xiao Qi (forthcoming); the following two paragraphs are summaries of a part of my arguments in this paper.

3.3 The Objects and Scope of Moral Obligation

than transient actions to a considerable degree. Hume's argument which rules out actions as proper objects of moral sentiments, then, should also rule out motives as such proper objects, and motives should be ascribed with the qualities of virtuousness or viciousness only insofar as they express some durable traits of character.

When a character trait is either constituted or signified by a specific motivating passion⁸¹, of course, there is no significant difference between saying that this character trait is virtuous and saying that the motive which signifies it is virtuous. The virtues (which are also duties) that Hume uses as examples for the Motive Principle, namely, gratitude and parental love, both fall in this category: these character traits can be characterised as one's tendency to feel the motives of gratitude or parental love. When we turn to Hume's catalogue of virtues provided in Book 3, part 3 of the *Treatise* or the second *Enquiry*, however, we may observe that not all qualities there are characterised by specific motives. One group of noteworthy exceptions is a number of traits that would fall in the category of intellectual virtues in the Aristotelian tradition, though Hume does not adopt the distinction between moral and intellectual virtues in his own taxonomy. Such virtues, like "prudence, penetration, discernment, discretion" (EPM App.4.2; 313), are described as "a disposition or turn of mind" (EPM 6.29; 246) or a "character, or peculiar understanding" (T 3.3.4.6; 610).⁸² They can have a significant influence on one's actions (EPM App.4.2; 313) and thereby be expressed by actions, but they do so by enabling one to make certain causal inferences or instantiate certain beliefs, and the latter process, as Hume argues, cannot motivate by itself (T 2.3.3.2-3; 413-14). Another group of exceptions are qualities in Hume's catalogue which, like cleanliness and decency, are principally specified by the agent's disposition to take on certain types of external appearances. One may have various reasons to cultivate such qualities, but we usually do not investigate these motives when ascribing these qualities to people.⁸³ Finally, qualities like the strength of mind, which "implies the prevalence of the calm passions above the violent" (T 2.3.3.10; 418), are not characterised by particular types of passions but by structures between motivating passions. We praise a

⁸¹ The ontological status of character traits is a controversial topic among interpreters. Character traits are usually initially grasped as dispositional qualities of the mind (say, one is inclined to feel and act in such a manner under such types of circumstances), but it is controversial whether, on a more fundamental level, they should be reduced to non-dispositional qualities like particular perceptions. Jane McIntyre (1990, 199–201) and Eric Frykholm (2012, 87–92), for instance, maintain a reductionist reading, while Qu (2017, 646–55) holds a non-reductionist reading. I do not take a position in this debate here because both readings are compatible with my view that some traits of character are not constituted or signified by specific motives.

⁸² For discussions on Hume's view on the characterisation of intellectual virtues, see Baier (2008, 3–10) and Dan O'Brien (2012, 155–59).

⁸³ On this point, especially see Enrico Galvagni's (2022, 139–45) argument that the quality of decency consists merely of "a fair outside" instead of any internal qualities.

3.3 The Objects and Scope of Moral Obligation

person for their strength of mind because their ruling passion is calm but strong, but this passion does not need to be of a “virtuous” type; it can be morally neutral like self-love.⁸⁴ For all these virtues, when we praise one for having such qualities, we do not praise them for having any specific types of virtuous motives but for their taking specific types of means to achieve their ends, maintaining certain external appearances, or being able to follow stronger but less violent passions. It is implausible, then, that the Motive Principle is universally applicable to all virtues in Hume’s catalogue.

If the Motive Principle cannot be applied to all virtues, can it be applied to all duties? In other words, is it true, on Hume’s account, that “[w]hen we require any action ... we always suppose, that one in that situation shou’d be influenc’d by the proper motive of that action” (T 3.2.1.3; 477)? As I noted above, the examples Hume appeals to in the argument following this claim, namely, the duties of parental care and gratitude, are all motive-characterised qualities, and they are required because the motives which constitute or signify these qualities, namely, parental love and the passion of gratitude, are morally required. The same can also be said of other duties which Hume frequently talks about, like pity and humanity (T 3.2.5.6; 518; E-OC 33, 479). In the second *Enquiry*, however, when Hume discusses virtues which are useful to oneself, and when he argues against the distinction between virtues and talents, he appeals to the notion of “duty to oneself” (EPM 6.13, App.4.21; 238, 322). In both contexts, he argues that the lack of some privately useful virtues leads to self-disapprobation and humility upon self-reflection. Not only the approbation we give to virtues and talents are made on the same footing, but also the disapprobation we give to one’s lack of them: “[a] blemish, a fault, a vice, a crime; these expressions seem to denote different degrees of censure and disapprobation; which are ... at the bottom, pretty nearly of the same kind or species” (EPM App.4.21; 322); for instance, one is usually “deeply mortified” when reflecting on one’s folly or impudent conduct (EPM App.4.3; 314-15). If many ‘talents’ are not praised for involving any virtuous motives, and if the lack of them could generate a sentiment of moral obligation just like the lack of gratitude and parental love, it might follow that some qualities, which are regarded as one’s duties, are not required due to any obligatory motive. Whether this observation can actually work to shed doubt on the Motive Principle’s applicability to all duties still depends on Hume’s view on the scope of duties, but this observation is sufficient to show that one cannot

⁸⁴ On this point, especially see the discussions of McIntyre (1990, 201) and Radcliffe (2018, 160–82).

3.3 The Objects and Scope of Moral Obligation

presume, based on Hume's examples of parental love and gratitude, that all duties are motive-characterised.

The question about the object of moral obligation, then, leads us to the question about the scope of duties, especially about whether the scope of natural or non-artificial duties extends beyond Hume's usual examples which are roughly different expressions of benevolent motives. As I noted above in Section 3.1, Hume does not provide a productive definition of duty like his four-branched definition of virtue, and if the sense of duty arises as the disapprobation of a vice or the absence of a virtue, Hume does not give any determinate account about whether there is a duty to avoid *all* vices or to have *all* virtues. I argued in Section 3.2 that the common course of human nature may provide a standard for duties by guiding people's expectations for virtues or remedies for vices. This standard can explain why many 'natural' duties that Hume discusses are characterised by instinctive motives like kindness to children: such motives are regarded as morally obligatory not only because the absence of them is harmful but also because they are generally expected for an average virtuous person. The same standard, meanwhile, cannot rule out the duties to have certain 'abilities' or 'talents' which are not characterised by motives, since a certain degree of such qualities can also be generally expected in a community, and the lack of them is also usually detrimental to one's own or others' interests. Does Hume endorse, then, any other guidelines for judgements about duties?

Let me consider a group of suggestions made in the literature. Shaver (1992, 555) argues that for Hume, "one has a duty or an obligation to avoid any vice", and Brown (1994, 22–31) argues that the sense of duty arises "when our self-conception clashes with the ideal of character", an ideal of a "fully loveable" person that one constructs based on experiences about what others regard as loveable and morally meritorious, which indicates that one has a duty to have every virtue which can be expected for oneself. Especially in Brown's exemplarist reading, it would be morally obligatory for one not only to love one's family, feel pity for the miserable, and abide by the rules of justice, but also to have virtues like wisdom, application, wit, and tranquillity of soul – all the virtues a character ideal like Hume's example of Cleanthes (EPM 9.2; 269–70) has. Baier (2010, 132–33) argues against such readings based on Hume's pessimistic position concerning individual agents' ability to change their characters. Many virtues on Hume's list, especially those which are more traditionally regarded as 'natural abilities' or 'talents', cannot be acquired voluntarily, and Hume remarks on this point that "it being almost impossible for the mind to change itself in any considerable article, or cure itself of a passionate or splenetic temper when they are natural to it" (T 3.3.4.3; 608). If one cannot

3.3 The Objects and Scope of Moral Obligation

acquire a virtuous or avoid a vicious character trait voluntarily, Baier argues, “it would be inhumane in the extreme to say that there is an obligation” to acquire or avoid them. Based on the same reason, Baier suggests that Hume only acknowledges obligations to observe artificial moral rules because it is possible for one to redirect one’s passions with the help of society-wide cooperative institutions. Besides artificial obligations, Baier reads Hume as using the concept ‘duty’ for “a fairly definite moral requirement on action arising out of some (possibly natural) station a person occupies”, like the duties of a parent, friend, wife, or citizen, and the scope of duty cannot be extended to role-independent qualities like cheerfulness and patience (2010, 135).

I argued above that I do not agree with Baier’s interpretation that Hume draws a clear-cut distinction between the notions of duty and obligation. The two restrictive conditions that Baier suggests, namely, (i) voluntary acquisition and (ii) social roles or stations, however, are both worth considering. The first condition is an application of the principle ‘ought implies can’. I am sceptical, nevertheless, about whether this principle can be ascribed to Hume’s conception of duty and obligation, because it appears that this contradicts his explicit view that the duties of parental love and gratitude are about motives rather than external actions. As I observed above, in the case of the ungrateful man, he can never fully discharge his moral obligation and get rid of his self-disapprobation by merely performing grateful actions since what is *ultimately* obligatory is the grateful motive. This motivating passion, or one’s tendency to feel this passion under suitable circumstances, cannot be raised at will. Though it is possible for one to interfere with the dynamics of one’s passions by reflections and other techniques, the force of such interference is limited, and a considerable part of one’s character traits is the consequence of luck (with respect to both one’s ‘natural temper’ and the influence of custom and childhood education) rather than voluntary choice.⁸⁵ In the second *Enquiry*, Hume remarks on those people who are “deficient in their duty to society” that their characters are morally deficient not “because they would not *wish* to be generous, friendly, and humane; but because they do not *feel* themselves such” (EPM 9.21; 282, my italics). Since *some* qualities, which Hume explicitly acknowledges as duties, cannot be voluntarily acquired, there is reason to doubt whether Hume might endorse voluntary acquisition, or the principle ‘ought implies can’, as a

⁸⁵ Hume’s pessimism about one’s ability to change one’s character is most explicitly addressed in his essay “The Sceptic” (E-Sc 28-55; 168-80), where he acknowledges the force of study, habit, and philosophical reflections in refining one’s temper but insists that such changes are not always in one’s power. For a helpful summary of scholarly controversies on Hume’s views on the possibility of character change and the approaches for character development, see the discussion of Ruth Boeker (2022, 124–27).

3.3 The Objects and Scope of Moral Obligation

standard for judgements about duties. Hume's position on this point may not be too demanding as he takes the common course of human nature as a guide for the sense of duty, and in this respect, it may be said that he endorses 'ought implies can' on the level of the general condition of mankind, but not of individuals. When one is so unfortunate as to, by one's natural temper, lack a virtuous instinct common in human nature, it has to be acknowledged that one can never, probably, genuinely fulfil one's duty in this respect.

Baier's suggestion that duties have a more limited scope than virtues because they arise from specific roles and stations that one occupies in a society can be more plausibly ascribed to Hume. I agree with Erin Frykholm (2016, 616–26) that for Hume, moral evaluation is, in general, sensitive to the evaluated agents' social roles, and this sensitivity is especially addressed in Hume's remark that when evaluating an agent's character, people confined their view to "the narrow circle, in which any person moves" and sympathize with those people in such a circle who are the usual or typical beneficiaries or victims of the agent's character traits (T 3.3.3.2; 602).⁸⁶ Whatever quality renders one "incommodious to those, who live and converse with him" is always regarded as "a fault or blemish, without any farther examination" (T 3.3.3.9; 606). Considerations about one's social roles, moreover, like those about the common course of human nature, guide people's general expectations about others' traits and actions; as Hume remarks in the second *Enquiry*, expectations about the degree of good that a person brings about are always "[i]n proportion to the station which a man possesses, according to the relations in which he is placed", and disappointment in such cases induces blame (EPM 5.38n; 225). As in Hume's view, the consideration that "there be no relation of life, in which I cou'd not wish to stand to a particular person" also provides "the ultimate test of merit and virtue" (T 3.3.3.9; 606), however, there is no strong textual support for the reading that social roles may draw a distinction between the virtuous qualities which are also duties and those which are not.

We have to conclude here, therefore, that Hume does not provide any determinate account of the scope of duties in his works and whether it has any significant difference from the scope of virtues. It is probable that he does not do so because of a lack of interest: as I observed in

⁸⁶ While Frykholm holds that the 'narrow circle' plays an important role in Hume's theory of moral evaluation, Baier (2006, 113–14) argues that considerations about one's 'narrow circle' only concern the evaluation of benevolence. I agree with Frykholm that Baier's restrictive reading is unnecessary as the 'narrow circle' can be read as a shorthand for those people associated with the evaluated agent and typically influenced by the latter, and who fall into such a 'circle' must be determined by the specific agent and character trait under concern (2016; 622–23). On this point, also see the discussion of Nir Ben-Moshe (2020, 443–44) with which I largely agree.

3.3 The Objects and Scope of Moral Obligation

Chapter 1, Hume did not intend to compose his moral works as a system of duties and obligations. He did not reject using the language of duty, including some traditional categories of duties, like the duties to oneself and to others, in his virtue-centred moral works, but his accounts of duties, especially when they concern *natural* or non-artificial virtues, are usually made in passing, and it can hardly be expected that such remarks could make up a systematic theory. Also, Hume might not find a strong need to clarify the relationship between virtues and duties as he did not mean his moral works to provide guidance for people's moral practices, for which a catalogue of duties would be helpful. As for his descriptive and explanatory interests, he might find the accounts of the features of the sense of duty as sufficient for capturing people's feelings, and since many virtues are actually also regarded as duties, and since many of his readers would not find the switch between the languages of virtue and duty problematic and in need of extra explanation⁸⁷, he might think that giving a full account of duties besides his accounts of virtues seems redundant. This is consistent with my reading that judgements about virtues and duties have different types of sentimental bases, and even if, in Hume's view, the scopes of virtue and duty are *coextensive*, their different sentimental bases (that is, the approbation of certain qualities and disapprobation of the lack of them) will still make the same mental quality virtuous and morally obligatory for different reasons and in different degrees. A common degree of gratitude, for instance, may not be regarded as significantly virtuous, but the lack of it causes strong moral disapprobation; an exceptional degree of wisdom, fortitude, *or* benevolence, which exceeds usual expectations, at the same time, may hardly be taken as duties for most people. Hume's account accommodates all such common-sense opinions well, but concerning non-artificial virtues or duties, maybe it does not go further than this.

Meanwhile, another reason why the texts I have examined in this chapter cannot provide a full account of the scope of duties is that, given Hume's view that the *idea* of obligation arises from the conventions of justice (T 3.2.2.11; 490-91), it is probable that no full account of the origin of *all* judgements in terms of duty and obligation, *including* the duties to have certain natural virtues, can be given without taking justice and its conventional origin into concern. As his account of the origin of moral obligation is yet to be completed, the scope of moral obligation may remain to be determined. It is probable, then, to have a better understanding of

⁸⁷ As Haakonssen (2012, 276) observes, under the influence of Stoicism and natural law theories, many Scottish moral authors in the mid-eighteenth century used the languages of virtue and duty in an interchangeable manner. Hume's mixed usages of these two languages, especially in the second *Enquiry*, may be read as influenced by such a context or as the result of his attempt to address his thoughts to such a group of readers, although this does not suggest that Hume shared his contemporaries' approaches to integrating ancient and modern moral thoughts.

3.4 Conclusion

Hume's view on 'natural' or non-artificial duties, we need to turn our attention to the artificial ones, which I will examine in the following chapters.

3.4 Conclusion

This chapter was an attempt to delineate the basic framework of Hume's conception of moral obligation. Starting with Hume's rudimentary delineation of moral obligation in the *Treatise*, I argued that the sentiment of moral obligation is an unpleasant sentiment caused by ideas of the absence of a mental quality or the omission of an action. In alignment with his sentimentalist approach to defining virtue, duty, or the object of moral obligation, may be defined as 'whatever action or quality of the mind the neglect, or non-performance of which gives to a spectator the displeasing sentiment of disapprobation, or a sentiment of moral obligation'. I suggested that the sentiment of moral obligation should be understood as a derivative form of moral disapprobation, which arises when the absence of a virtuous quality is regarded as vicious, or the disapprobation of a vicious quality is associated with ideas of remedies for this vice. People's general expectations about an averagely virtuous character may play an important role in the formation of this sentiment through both possible routes of generation, and this explains why Hume claims that the sense of duty is always guided by the common course of human nature. Regarding the object of moral obligation, I observed that in Hume's view, when the sense of duty is felt about the absence of a virtuous quality, merely performing external actions which signify such a quality cannot make one fulfil one's duty since in such cases, the ultimate object of moral obligation is the internal quality. In some cases, this suggests that one ought to be motivated by the proper passion which is virtuous and generally expected, and in other cases, this means that one ought to have certain abilities or habits which enable one to act in a certain manner. This means that Hume does not endorse the principle 'ought implies can' on the level of individual agents, at least when some 'natural' duties are under concern. These discussions, finally, cannot provide a full account of the origin and scope of moral obligation in Hume's view, partly because it is probable that all judgements about obligation are influenced by the convention of justice. From the next chapter on, I will investigate why Hume pays much more attention to the topic of obligation in his theory of justice compared to his account of the natural virtues, and how he develops a story of the conventional origin of justice to explain some special features of the obligation to justice.

Chapter 4

Moral Obligation and the Artificiality of Justice

I explored the basic framework of Hume's conception of moral obligation in the last chapter. According to Hume, we judge what is morally obligatory by our moral sentiments, and the sentiment which provides us with such ideas is an unpleasant feeling caused by ideas of the absence of certain actions or mental qualities. This sentiment differs from the approbation of virtue and the disapprobation of vice but can be derived from them. The typical objects of the sentiment of moral obligation, or one's duties, are determined by the nature and causal origin of this sentiment. When this sentiment arises upon one's failure to meet a certain expectation for virtue, what is regarded as one's duty is to have this expected virtuous *quality*, and actions are morally obligatory only insofar as they express a quality which typically causes them.

In some cases, however, the characteristics of the sentiment of moral obligation cannot be fully accounted for with this basic model of approbation, expectation, and humility caused by the moral deformity in failing to live up to character expectations. This is the sentiment of moral obligation we feel about observing the laws of justice, and it falls outside of this basic model because the way we feel this sentiment is in tension with the inclination of moral sentiments in general. I argue in this chapter that it is in light of this tension that Hume argues in the *Treatise* that justice is an 'artificial virtue'. As I observed in Chapter 1, there is an interesting imbalance between Hume's interest in the topic of moral obligation in his theories of natural and artificial virtues. He rarely talks about these notions when he discusses moral sentiments in general or the natural virtues in particular, but he uses them frequently in his theory of justice and allegiance. This can be explained, as I argue in this chapter, by the special conspicuousness of moral obligation in our judgements about justice. Because it deviates from how people usually approve of other virtues and require for virtue-derived duties, there is a special need to explain the origin of this sentiment.

This chapter focuses on how Hume's observation about the moral obligation to justice leads him to a theory of artificial virtues by examining his arguments on the difference between the natural virtues and justice, or his 'Artificiality Arguments' as I shall call them. Hume offers five such arguments in Book 3, part 2 of the *Treatise*: one opening argument in *Treatise* 3.2.1,

which he repeats in a more abridged form in 3.2.5; one argument which specifically pertains to promises in 3.2.5; and three arguments in his “father reflections concerning justice and injustice” in 3.2.6. I start in Section 4.1 to examine *Treatise* 3.2.6 in which Hume discusses how the sentiment of moral obligation to justice diverges from the usual ways we feel about virtues and vices: it raises inflexible requirements to observe a set of universal laws; it attends ideas of specific actions or objects, and it leads to judgements which do not allow for degrees. In Section 4.2, I paraphrase the argument in *Treatise* 3.2.1 which focuses on people’s motives to perform just actions and introduce scholarly controversies on it. I then argue in Section 4.3 that, unlike the interpretation of many scholars, this argument does not set up a search for a missing motive apart from the sense of duty which may characterise the virtue of justice. On the contrary, the gist of this argument does not differ from those in *Treatise* 3.2.6, namely, it highlights a special feature of the sentiments which we feel about just actions which is in tension with the natural inclination of the moral sentiments.

4.1 Obligation, Virtue, and Law: *Treatise* 3.2.6

The first six sections of Book 3, part 2 of the *Treatise* make up a partly self-contained unit in Hume’s moral theory. It starts with an argument which proves that the sense of justice arises from conventions. Then, there are four sections which explain its causal origin: two on property, one on the transference of property by consent, and one on promises. *Treatise* 3.2.6 concludes these accounts with three new arguments to prove that these “three fundamental laws of nature” “are entirely artificial, and of human invention” (T 3.2.6.1; 526). Compared to *Treatise* 3.2.1, Hume’s arguments in *Treatise* 3.2.6 have attracted less scholarly attention. The latter arguments are nevertheless more informative: unlike *Treatise* 3.2.1, which precedes Hume’s own positive theory of justice and is constructed with a group of suppositions that he appears to take for granted without sufficient explanation, *Treatise* 3.2.6 reflects on what *has been* explicated on the characteristics of the laws of justice. Accordingly, I discuss *Treatise* 3.2.6 before 3.2.1. As Hume means his arguments in *Treatise* 3.2.6 to prove the artificiality of justice, he should not presume his account of the conventional origin of justice in them, and I will bypass the related comments in my examination.

The moral obligation to justice plays a crucial role in all of Hume’s three arguments in *Treatise* 3.2.6. He starts with an argument on how the sense of duty contributes to the quality of things which we call ‘property’. This quality does not consist of an “external and corporeal relation” between a person and an object, like occupation, because the same relations can be

observed between “brute creatures” and objects but do not make up any relation of property. Instead, property consists of an “internal relation” which a person-object relation causes in people’s minds through their moral sentiments:

Now ’tis evident, this external relation causes nothing in external objects, and has only an influence on the mind, by giving us a sense of duty in abstaining from that object, and in restoring it to the first possessor. These actions are properly what we call *justice*; and consequently ’tis on that virtue that the nature of property depends, and not the virtue on the property. (T 3.2.6.3; 527)

Only when people react to an external relation between a person and an object with a sense of duty to abstain from this object, that is, only when an ‘internal relation’ is established between this external relation and people’s moral sentiments, is this object regarded as this person’s property. In both this section and earlier, Hume refers to “property, and right, and obligation” as “*moral relations*” (T 3.2.6.8; 531, also see T 3.2.2.11, 3.2.3.10n4; 490-91, 510), and when he uses this phrase, he always talks about these three notions in a group. By ‘moral relation’ here, he should mean the tri-partite ‘internal relation’ between a person, an object, and people’s general sentimental reaction. It differs from the relations of ideas which Hume criticizes in his arguments against moral rationalists because the sense of duty is its essential constitutive part, which alone makes the person-object relation morally significant. If the sense of duty ceases to work, the quality of property will soon vanish (T 3.2.6.3; 527).

Based on this observation, Hume argues that justice cannot be a natural virtue. We call the actions to abstain from others’ property and restore others’ property to its owner actions of ‘justice’, and ‘property’ derives its meaning from the sense of duty that we feel about abstaining from it. If justice was a natural virtue, Hume argues, we would not rely on the sense of duty to specify what is just or unjust as there would be some ‘*natural*’ quality in person-object relations which can explain the moral value of such actions. There is, meanwhile, no such natural quality: actions are only specified as ‘just’ because they pertain to someone’s ‘property’, not because they pertain to some objects which, for instance, one first occupies or possesses for a long time. By a ‘natural quality’ here, Hume should mean a ‘non-moral’ quality or a quality different from the type of moral relation with which he defines property. Hume clarifies in *Treatise* 3.1.2 that he uses the term ‘natural’ sometimes as opposed to ‘artificial’ and sometimes as opposed to ‘moral’ (T 3.1.2.9; 475-76). Now, Hume argues that justice is not natural (*non-artificial*) by pointing out that property is not a natural (*non-moral*) quality, but this inference seems unwarranted. Presumably, if property was an ‘artificial’ moral quality, the virtue of abstaining from others’ property would also be artificial, but Hume does not argue why the *moral relations*

of right and obligation must be artificial or convention dependent. This argument, therefore, by itself, does not establish that property is an artificial moral quality. So much for Hume's first argument.

The topic of moral relation continues in Hume's second argument. This argument begins with Hume's observation that:

[A]ll kinds of vice and virtue run insensibly into each other, and may approach by such imperceptible degrees as will make it difficult ... to determine when the one ends, and the other begins ... 'tis certain, that rights, and obligations, and property, admit of no such insensible gradation, but that a man either has a full and perfect property, or none at all; and is either entirely oblig'd to perform any action, or lies under no manner of obligation. (T 3.2.6.7; 529)

As a second step, Hume argues that because "rights, and obligations, and property" "depend entirely upon justice and injustice", if the former qualities are not susceptible to insensible gradation, neither should the latter (T 3.2.6.7; 530). Finally, because "vice and virtue, moral good and evil, and indeed all *natural* qualities" allow for degrees, justice and injustice cannot be '*naturally* either vicious or virtuous' (T 3.2.6.7; 530).

When Hume observes that obligation does not admit of gradation, he talks about it as a constitutive part of the "moral relations" which makes an object one's property or makes an action obligatory for a person (T 3.2.6.8; 531). Such relations are unsusceptible of gradation since "[a]n object must either be in the possession of one person or another. An action must either be perform'd or not" (T 3.2.6.8; 530). On this point, Jonathan Harrison (1981, 158–59) argues that Hume cannot derive from this observation his conclusion that justice is artificial because he is "guilty of a cross division": nature and artifice on the one hand, and the moral evaluation of actions and persons on the other. When we talk about whether an *action* is obligatory or right, Harrison argues, the questions are always to be answered by either/or no matter whether it concerns a natural or artificial virtue. On the other hand, the evaluation of a person's character always admits of degrees.⁸⁸ What Hume should distinguish here is virtue and obligation, not nature and artifice.

I agree with Harrison that the difference between the evaluation of actions and character is of crucial importance here, but I disagree with his comments on the natural virtues. Let me

⁸⁸ For a similar suggestion, see Árdal (1966, 187). Baier (2010, 56–61) argues that Hume can be free of Harrison's charge if Hume raises a very high standard of the virtue of justice, namely, one is a just person if one never breaches the rules of justice. This, however, Baier comments, contradicts our common life experience in which one can be more or less just.

use Hume's favourite example of a natural virtue, that is, benevolence, here. As I argued in Chapter 3, the sentiment of moral obligation, in this case, is derived from the approbation of the virtue of benevolence, and an action is regarded as morally obligatory only insofar as it is typically caused by a benevolent motive. On this account, when evaluating an action in this aspect, one needs to first consider the moral quality of the trait of character that this action typically expresses. One needs to consider what a benevolent person usually does in such and such circumstances when motivated by benevolence. Evaluations made by such an inference are usually not precise and are susceptible to change according to circumstances, as Hume argues in his third argument in this section (T 3.2.6.9; 531-32). Also, as a quality can be more or less virtuous or obligatory due to its specific psychological makeup, its typical consequences, and people's expectations, such variation will also apply to the evaluation of actions as long as the latter is derived from the former.

Judgements on moral obligation which pertain to actions about property and promises, on the other hand, do not admit of degrees because they are always precise: an action is either fully obligatory in such respect or not obligatory at all. This differs from how we evaluate the moral goodness of virtues like benevolence which Hume calls a 'natural quality'. The 'natural' here means 'non-artificial' rather than 'non-moral' since Hume infers from it that justice is not 'naturally virtuous', and it is self-contradictory to claim that something is 'non-morally virtuous'. This point explains the underlying reason for Hume's first argument in which he infers from the premise that property is not a natural (non-moral) quality to the conclusion that justice is not a natural (non-artificial) virtue. Taken by his second argument, the reason which supports this move should be that the *moral* quality of property is *artificial* because the sense of duty which constitutes this quality is felt in a way which differs from how we naturally (non-artificially) evaluate virtues and virtue-derived duties. The distinction between the moral evaluation of actions and character traits, then, should be explained by the distinction between nature and artifice. The difficulty which remains in this argument is that Hume does not clarify what he means by the 'justice' on which the obligation to abstain from others' property "entirely depend" (T 3.2.6.7; 530). If it refers to a virtuous character trait but does not admit of degrees, this contradicts Hume's claim that all virtues and vices admit of degrees. If it does not refer to a virtue, then it seems that the conclusion of this argument should be that justice is not a *virtue*, not that it is not a *natural* virtue.

As Hume's second argument helps to explain his first argument, his third argument helps to explain the second one by bringing to the fore the relationship between the moral relations

of ‘right, obligation, and property’ and a set of universal and inflexible moral rules. Here, Hume explicates how the “common principles of human nature” are susceptible to variation as influenced by particular circumstances, and accordingly, when moral evaluation follows such natural (non-artificial) principles, “the morality must be susceptible of all the same variations, which are natural to the passion” (T 3.2.6.9; 532). It then turns out that the contrast between whether certain moral qualities admit of degrees is only one aspect of the differences between moral judgements based on the rules of justice and those based on the “ordinary course of our passions and actions” (T 3.2.6.9; 532). As the features of such rules contradict the natural (non-artificial) characteristics of moral evaluation, the rules of justice cannot be natural.

Hume’s three arguments in this section, then, can be read as a complete line of reasoning which starts with the definition of property and moves on to explain the underlying reasons for this phenomenon. Their conclusions can be summarized as follows:

- (1) Because property is not a natural (*non-moral*) quality, justice, which is specified in terms of property relations, is not a natural (*non-artificial*) virtue.
- (2) The *moral* qualities of property, right, and obligation deviate from the natural (*non-artificial*) moral qualities of virtue and vice as they do not admit of degrees.
- (3) The rules of justice deviate from the natural (*non-artificial*) principles of moral evaluation as they are universal and inflexible.

(3) can explain (2) since Hume argues that “property, right, and obligation are determin’d” by the rules of justice (T 3.2.6.6; 528), and (2) can explain (1) since it explains why the moral quality of property is artificial. This is consonant with Hume’s claim of his target in this section at its beginning: because “these laws, however necessary, are entirely artificial, and of human invention”, “*consequently* ... justice is an artificial, and not a natural virtue” (T 3.2.6.1; 526). The virtue of justice is artificial because it is specified in terms of moral relations determined by artificial moral rules.

This line of reasoning appears clear and tidy, but the difficulties mentioned above remain. One implication of this line of reasoning is that ‘justice’ is, at least fundamentally, a quality of actions determined by a set of laws, and the virtue of justice depends on the laws of justice and its corresponding obligation to be defined. Hume, at the same time, however, repeats in all three arguments that property and obligation depend on the *virtue* of justice. It may appear that he should have better dropped this problematic claim to make his arguments more tenable, as

commented by many scholars. Darwall (1995, 296–98, 312–14), for example, argues that the rigid obligation to justice should be a type of “rule obligation” which differs from moral obligation which consists in the moral approbation and disapprobation of character, and Harris (2010a, 27–28, 35–36) argues that the obligation Hume addresses here corresponds to perfect obligation in modern natural law theories which requires for external actions or inactions instead of any virtuous motivational states. Hume could have reason to reject such easier answers as for him the evaluation of action has to be derived from the evaluation of character, and also since he does not embrace a conception of moral laws as the product of the will of a lawgiver or a rational order in the nature of things and actions. If Hume’s theory of convention can account for the features of the moral obligation to justice, it may be hoped that this theory can also account for his shift between virtue and law in these arguments. Only after a proper discussion of his theory of the conventional origin of justice in the next two chapters can we return to this subject.

In his third argument, the claim that property depends on justice leads Hume to a further remark that:

[P]roperty is perfectly unintelligible without first supposing justice and injustice; and that these moral qualities are as unintelligible, unless we have motives, independent of the morality, to impel us to just actions, and deter us from unjust ones. Let those motives, therefore, be what they will, they must accommodate themselves to circumstances, and must admit of all the variations, which human affairs, in their incessant revolutions, are susceptible of. They are consequently a very improper foundation for such rigid inflexible rules as the laws of nature; and ’tis evident these laws can only be deriv’d from human conventions. (T 3.2.6.10; 533)

By ‘justice’ in the first statement, Hume should mean the *virtue* of justice since if he means the laws of justice or qualities of just actions, justice does not need to depend on any motives to be intelligible. Hume does not explain, however, why a notion of the law of justice, independently of the virtue of justice, is not sufficient to make property intelligible. With this step set aside, there are still three problems in this argument. First, Hume does not explain why the quality of justice is only intelligible if we have motives “independent of the morality” to perform just actions. Second, according to Hume’s third argument, *non-artificial* motives are prone to variations, but in this paragraph, Hume ascribes this feature to all *non-moral* motives and does not explain why he makes this change. Third, if Hume was seriously committed to the first two claims in this passage, it would follow that both property and justice were unintelligible, not that they were artificial or derived from human conventions, and Hume does not explain how

this artificial origin could help to solve the problem regarding the intelligibility of property and justice.

All these difficulties can also be observed in Hume's longer and notoriously controversial argument in *Treatise* 3.2.1, which I will turn to in the next section. For a brief overview of my reading, I will argue that Hume's second move in the passage above, that is, the intelligibility of the virtue of justice depends on a non-moral motive, is proposed under a *false* assumption that justice is a natural virtue which can be approved in a fictional 'state of nature'. This explains both Hume's shift from non-moral to non-artificial motive in his third step and his conclusion concerning the artificiality of justice. Now let me turn to *Treatise* 3.2.1.

4.2 Obligation, Virtue, and Motive: *Treatise* 3.2.1

Hume starts his arguments on the artificiality of justice in *Treatise* 3.2.6 with the definition of property. The sentiment of moral obligation plays a central role here: it makes the relation between a person and an object a 'moral relation', and property and right, taken as moral qualities of this object and its owner, in this sense, depend on this sentiment being intelligible and having moral significance. The way we perceive these moral qualities differs from the way we approve of virtues and disapprove of vices, as Hume argues, as the former follows a set of universal and inflexible rules. Here arises a tension between our moral judgements about justice, which are formulated in terms of rights and obligations, and the natural tendency of the moral sentiments which focus on persons and characters. Instead of distinguishing two types of moral evaluations, Hume argues that here lies the difference between artificial and natural 'virtues', though we do not know how the artificial origin of justice helps to resolve this tension.

Hume's Artificiality Argument in *Treatise* 3.2.1, meanwhile, focuses on people's motives to perform just actions. According to one popular reading of this section, Hume argues that justice is an artificial virtue by demonstrating that it needs to be characterised by an artificial virtuous motive other than the sense of duty. In my reading, however, this argument is not meant to start up an inquiry for such a motive. Like *Treatise* 3.2.6, this argument highlights a special feature of the sentiment of moral obligation to justice and concludes that this *sentiment* is artificial. I paraphrase Hume's argument, locate several problems in it, and introduce current interpretations in this section, and I propose my own reading in the next.

The argument in *Treatise* 3.2.1 has three parts. The first part begins with two observations which lead to the principle that "all virtuous actions derive their merit only from virtuous

motives” (T 3.2.1.4; 478), which I call the ‘Motive Principle’ in the last chapter. As I noted above, many scholars read this principle as universally applicable to all types of moral evaluations of actions in Hume’s theory: for all virtues, vices, and duties expressed by actions, the ‘ultimate objects’ of moral evaluation are motives. I argued in the last chapter that it is unlikely that this is true: at least some virtues in Hume’s catalogue are not praised for any virtuous motivating passions which constitute or signify virtuous traits of character, and it remains to be explained why Hume starts this argument with this proposition *as if* it was a universal principle. This is the first problem with this argument, but let me first leave it here. After putting forward the Motive Principle, Hume keeps discussing virtue and duty on the same footing. Since he draws this principle from two observations, one on how we *praise* actions and the other on how we *require* actions, I do not take this argument as a sign that Hume finds no distinction between the sentiment of moral obligation and the approbation of virtue. I will follow Hume in using ‘virtue’ as a representative for ‘virtue, vice, and duty’ in my paraphrase of this argument.

In the first part of the Artificiality Argument in *Treatise* 3.2.1, Hume argues that:

From this principle I conclude, that the first virtuous motive, which bestows a merit on any action, can never be a regard to the virtue of that action, but must be some other natural motive or principle. To suppose, that the mere regard to the virtue of the action, may be the first motive, which produc’d the action, and render’d it virtuous, is to reason in a circle. (T 3.2.1.4; 478)

Let me refer to this passage as Hume’s ‘Circle Argument’. Because, according to the Motive Principle, the ultimate objects of moral evaluation are motives, any moral value ascribed to actions must be grounded on and explained by the value of the type of motives which typically cause such actions. A “regard to the virtue of an action”, which is either constituted by or derived from one’s approbation of this action⁸⁹, can work as a virtuous motive, but it needs a *precedent* virtuous motive to provide the action at issue with its moral merit. At the bottom of such a reductive process, there must be a virtuous motive which can work independently of moral sentiments. Otherwise, an infinite regress would occur, and the initial regard to virtue would be groundless. It follows from this line of reasoning that by ‘natural motive’ in the first sentence of this passage, Hume means ‘*non-moral* motive’.

⁸⁹ I use “either constituted by or derived from” to keep neutral with the controversy over whether Humean moral sentiments can motivate by themselves or only with the help of other passions, as I noted in Chapter 1.

After this short argument, Hume offers two examples to illustrate his conclusion: the duty of parental care and the virtue of humanity. In both cases, what is more fundamentally morally obligatory or virtuous are the motives of parental love and benevolence. Then, he raises an ‘undoubted maxim’ that:

[N]o action can be virtuous, or morally good, unless there be in human nature some motive to produce it, distinct from the sense of its morality. (T 3.2.1.7; 479)

Here arises the second problem with this argument, that is, how to understand the restrictive condition “*in human nature*” in this ‘maxim’. Noteworthily, in the Circle Argument, Hume does *not* suggest that the non-moral virtuous motive required there should be ‘natural’, taken in the sense of being instinctive or original to human nature. Don Garrett (2007, 261, 271–72; 2014, 264–68) suggests that this ‘undoubted maxim’ should be distinguished from the conclusion of the Circle Argument, as the former requires a natural (*non-artificial*) motive, and it does not say that this motive needs to be virtuous. Though this distinction could make sense based on Hume’s literal expression, it is not clear how Hume ‘*derives*’, as proposed by Garrett, the ‘undoubted maxim’ from the conclusion of the Circle Argument, since Hume only puts between these two paragraphs two examples. Therefore, either the ‘undoubted maxim’ only repeats the conclusion of the Circle Argument, in which case ‘in human nature’ can be read in a weaker sense, namely, it should be possible to find such a motive in human minds, or Hume switches from the requirement for a non-moral motive to that for a non-artificial motive without sufficient explanation.

Before the second part of his argument, Hume discusses how one can be motivated by a sense of duty with the case of the ungrateful man who hates himself, which I have examined in Chapter 3. The second part of this argument begins with an imagined conversation. Hume asks why ‘I’, if someone has lent ‘me’ a sum of money, have reason or motive to repay this debt. For a person in his “civiliz’d state”, Hume remarks, it is “just and satisfactory” to answer this question by saying that “my regard to justice, and abhorrence of villainy and knavery, are sufficient reasons for me, if I have the least grain of honesty, or sense of duty and obligation” (T 3.2.1.9; 480). The motive specified here is a moral motive, and “great difficulty” will occur, Hume argues, if we try to find a non-moral motive to account for the moral merit of this action. To illustrate this difficulty, he considers three options for such a non-moral motive: self-love, extensive benevolence or regard to the public interest, and confined benevolence or regard to the happiness of a particular person or group of people. All options fail his examination. As a conclusion of this part, he argues that:

From all this it follows, that we have naturally no real or universal motive for observing the laws of equity, but the very equity and merit of that observance. (T 3.2.1.17; 483)⁹⁰

The “very equity and merit of that observance” refers to the regard for justice or sense of duty in the reply of the “honest man”. To interpret this conclusion properly, it is important to clarify what Hume means by a “real and universal motive for observing the laws of equity” or what he looks for in his search for such a motive. The meaning of “real and universal” can be inferred based on three clues in this argument. First, before examining the candidates, Hume repeats the Circle Argument and its conclusion, namely, “[a] virtuous motive, therefore, must precede the regard to the virtue” (T 3.2.1.9; 480): this posits the need for a *non-moral virtuous* motive which provides the ‘ultimate object’ of moral approbation and accounts for the moral merit of its consequent action.⁹¹ Second, the conditions such a “real and universal motive” should meet can be inferred from Hume’s eliminative process. Arguably, a ‘*real* motive’ is a motive which really exists. It rules out, for instance, an alleged “love of mankind, merely as such, independent of personal qualities” since based on Hume’s study of human nature, “there is no such passion in human minds” (T 3.2.1.11; 481). A ‘*universal* motive’ for justice should refer to a motive which can *constantly* produce just actions. Because, as suggested in the Circle Argument, actions only derive their moral value from the motives which typically cause them, a constant connection must be found between the virtuous motives and the actions under concern. Self-love is not such a constant motive since when it acts without constraint, it may lead to injustice and violence (T 3.2.1.10; 480). A regard to public interest is not such a motive since in some cases, like when a secret repayment is owed to a secret debt, the performance of this action is not related to the public interest by itself (T 3.2.1.11; 481). Private benevolence, finally, is not a universal motive since it fails to promote just actions on many occasions: when the debt is owed to a vicious man, a miser, a “profligate debauchee”, or when the borrower is in urgent need to provide for his family (T 3.2.1.13; 482). Third, as argued by Geoffrey Sayre-McCord (2016, 439), it follows from the Motive Principle that it is the virtuous motive which *specifies* which type of actions express this virtue. Only actions produced by this motive are to be considered virtuous *in this aspect* instead of others. Though self-love and benevolence can produce actions which conform with the laws of justice in certain cases, we cannot rely on

⁹⁰ The word ‘naturally’ was not in the published version of the *Treatise*; Hume added it manually to his personal copy of the *Treatise*, and P. H. Nidditch follows Hume in his revision of the Selby-Bigge edition.

⁹¹ Garrett (2007, 261–62), meanwhile, argues that this part of Hume’s argument searches for a motive which can satisfy the requirement raised by the ‘undoubted maxim’, which, in his reading, posits the need for a *non-artificial* motive which does not need to be virtuous. This is unlikely given how Hume starts his query in *Treatise* 3.2.1.9, *if* Hume means to separate the immediate conclusion of the Circle Argument and the ‘undoubted maxim’.

these motives to specify, in general, what actions are ‘just’. When excluding benevolence from the “original motive to justice”, Hume makes the interesting point that:

A rich man lies under a moral obligation to communicate to those in necessity a share of his superfluities. Were private benevolence the original motive to justice, a man wou’d not be oblig’d to leave others in the possession of more than he is oblig’d to give them. At least the difference wou’d be very inconsiderable. Men generally fix their affections more on what they are possess’d of, than on what they never enjoy’d: For this reason, it wou’d be greater cruelty to dispossess a man of any thing, than not to give it him. But who will assert, that this is the only foundation of justice? (T 3.2.1.14; 482)

Hume’s point here is that if both the moral obligation of generosity, which at least applies to rich people when someone in necessity turns to them for help, and the moral obligation to justice were specified by the same motive, i.e., benevolence, the requirements of these two duties would become *undistinguishable*, and two counter-intuitive consequences would follow. From the perspective of the rich person, she would only be obliged to respect the poor person’s property to the extent that her benevolence motivates her to give, and supposedly, where this benevolence terminates, she could feel free to deprive the poor person of his possessions. On the other hand, from the perspective of a spectator, as the rich person is more emotionally attached to her property than the poor person who has not yet received it, it might contradict the principle of benevolence to require the rich to help the poor. Both consequences contradict experience: there is a duty to leave people in peaceful enjoyment of their property regardless of the benevolent affections one feel about them, and there is a duty for the rich to share their property to help the poor. Confounding the grounds on which people evaluate these two duties would threaten their recognition of both. To save both, one needs to specify these duties’ requirements in different ways, and if the contents of these duties are specified by some original motives, their corresponding motives should be distinct. A “real and universal motive for observing the laws of justice”, then, as required by the conclusion of the Circle Argument, should be an existent motive which can constantly cause just actions and specify the content of the virtue or duty of justice. For short, I shall call it a ‘*characterising motive*’ given its function for specification.

Hume concludes this part of his argument that there is ‘naturally’ no characterising motive for justice other than the sense of duty. How to understand this conclusion, especially the ‘naturally’ that Hume inserted in his manuscript revision of the *Treatise*, is the third problem which haunts the interpretation of this argument. Garrett (2007, 262) and Sayre-McCord (2016,

444) highlight Hume's addition and argue that this conclusion suggests that there is no *non-artificial* and non-moral characterising motive to justice.⁹² This is tenable in one sense but problematic in another. It is tenable because when Hume examines self-love and public benevolence, he argues that the '*natural* movements' of self-love usually cause injustice and that the "public interest is not naturally attach'd to the observation of the rules of justice" (T 3.2.1.11; 480). 'Natural' in both places means non-artificial. Meanwhile, there are two problems with this reading. First, I argued above that Hume sets out to search for a characterising motive to meet the requirement of the conclusion of the Circle Argument, and this conclusion states the need for a *virtuous non-moral* motive instead of a non-artificial motive. If Hume restricts his search to non-artificial motives, he should have clarified it somewhere, but he does not. Second, to recall, Hume phrases his conclusion as "we have naturally no real or universal motive for observing the laws of equity, but the very equity and merit of that observance". If 'naturally' here qualifies 'we have', it should apply to not only the negative part but also the positive part, making the statement here "we *naturally have* the very equity and merit of that observance as the real and universal motive to justice". The whole point of Hume's argument, however, is to prove that we do not naturally have a sense of justice: "the sense of justice ... arises artificially" (T 3.2.1.17; 483). Still, let me first leave this problem here.

In the third part of this argument, Hume derives his final conclusion from the conclusions of the former parts. As the sense of duty is the only "real and universal motive to observe the laws of justice", and as "no action can be equitable or meritorious, where it cannot arise from some separate motive" (i.e., a non-moral motive):

[T]here is here an evident sophistry and reasoning in a circle. Unless, therefore, we will allow, that nature has establish'd a sophistry, and render'd it necessary and unavoidable, we must allow, that the sense of justice and injustice is not deriv'd from nature, but arises artificially, tho' necessarily from education, and human conventions. (T 3.2.1.17; 483)

Here arises the fourth problem with this argument: what appears to follow from Hume's two previous conclusions is that the merit we ascribe to justice is groundless and that justice is not a genuine virtue or duty, not that justice is not a *natural* virtue. Let me now list the problems I have so far identified in this argument in a line to observe their connection:

⁹² Darwall (1995, 290) and Cohon (2008b, 166) do not find the word 'naturally' helpful and prefer to follow Hume's original wording to interpret this argument.

- (1) Hume starts the first part of this argument, namely, the Circle Argument, with the principle that all actions derive their moral value from their motives, but he does not endorse this principle as universally true in his theory of virtues. He does not explain why he relies so heavily on this principle in this argument.
- (2) It is unclear how Hume moves from the conclusion of the Circle Argument, which requires a virtuous non-moral motive to make actions virtuous, to an 'undoubted maxim' which requires a *natural* (non-artificial) non-moral motive.
- (3) It is unclear why Hume adds the word '*naturally*' to the conclusion of the second part of this argument, in which he searches for a virtuous non-moral motive which can account for the merit of just actions. This revision is also puzzling because we do not naturally have a moral motive for justice, given Hume's own theory.
- (4) It appears that this entire argument should have led to the conclusion that justice is not a virtue, not that justice is not a *natural* virtue.

Problems (2) and (3) are of similar structures: in both parts, Hume fails to clarify at what point the distinction between nature and artifice enters the story. It appears that in both parts, he starts with claims which are meant to be universally valid for all virtues and duties. At some point, however, he smuggles into his argument a further qualification, that is, 'natural' or non-artificial, to the subject issue. The same happens to his final conclusion. Fortunately, this suggests that if we can account for Hume's unclarified introduction of the natural-artificial distinction into this argument, we can account for the natural/artificial distinction in his final move.

Let me explain this connection with more details. Now, if both parts of this argument end up with universal conclusions, we will have:

C1: Actions can only be virtuous by deriving their merit from a virtuous non-moral characterising motive.

C2: There is no virtuous non-moral characterising motive for justice.

It should follow from them that:

C3: Just actions cannot be virtuous.

Hume's final conclusion, however, is that:

C3': Just actions cannot be naturally regarded as virtuous, or, in other words, they cannot be regarded as virtuous independently of conventions.

To reach this conclusion, at least *one* between C1 and C2 should be modified. Either, as argued by Garrett and Sayre-McCord, Hume's second conclusion is indeed:

C2': There is no non-artificial virtuous non-moral characterising motive for justice.

Or some alteration can be made to C1. It can be modified, as argued by Cohon (2008b, 171–83), if the Circle Argument is restricted to a subset of virtues, i.e., natural virtues. By showing that a scholium of this principle contradicts our conception of justice, Hume demonstrates that justice is not a natural virtue. Following Cohon's interpretation, we will have:

C1': For all natural virtues, actions can only be virtuous by deriving their merit from a virtuous non-moral characterising motive.

To make his argument for C3' valid, Hume can combine either C1 with C2' or C1' with C2. Alternatively, he can also put a 'natural' in both conclusions and suggest that the conclusion of the first part is:

C1'': For all *natural virtues*, actions can only be virtuous by deriving their merit from a *non-artificial* virtuous non-moral characterising motive.

This is the 'undoubted maxim' modified under the condition that the Circle Argument only applies to natural virtues. By combining C1'' and C2', Hume can also derive C3'. Finally, one can also suggest that Hume indeed believes that an artificial virtue is not a genuine virtue. We regard just actions as virtuous and morally obligatory due to some mistake in our conception of justice or our moral evaluation. There would be no need, then, to alter either C1 or C2 because Hume in fact acknowledges C3. This is the error theory interpretation suggested by Haakonssen (1981, 31–35) and Gauthier (1992, 421–22). I believe these four alternative readings cover most interpretations in current literature.

Choosing between these alternative readings is important at least to Hume's moral theory as delineated in the *Treatise* as this is how Hume first presents his distinction between natural and artificial virtues. How to understand this argument, as many scholars read it, determines how to construe both the 'naturalness' of natural virtues and the 'artificiality' of artificial virtues, and the latter aspect determines the task that this opening section leaves for Hume's following sections on the conventional origin of justice. If Hume's final conclusion is derived from C1

and C2', as argued by Garrett and Sayre-McCord, it will turn out that justice is artificial because its characterising motive is artificial, and the task for Hume's theory of convention is to reveal this artificial non-moral motive and to explain its causal origin. If, on the other hand, as argued by Cohon, Hume infers the artificiality of justice from C1' and C2, there will be no need to identify such a motive. What remains to be explained by Hume's theory of convention is why the sentiment of moral obligation to justice can arise on a ground other than people's motives so that *this* sentiment can serve as the characterising motive for justice without triggering a vicious circle. If Hume adopts C1'' and C2', either an artificial non-moral motive or an artificial moral motive can account for the artificiality of justice: this argument does not specify which it requires. Finally, if Hume endorses both C1 and C2 seriously, the artificiality of justice has to involve some cognitive or evaluative errors.

Despite its wide influence on the interpretation of Hume's theory of justice, how to make this choice is also quite controversial because it seems that each interpretative approach has its own difficulty. The first reading is confronted with challenges because it seems that in his account of the conventional origin of justice, Hume does *not* identify such a motive. Scholars who adopt this approach all attempt to find out this missing motive between lines in Hume's theory of convention: Garrett (2007, 272–76; 2014, 267) appeals to a “desire or standing disposition to govern or *regulate* one's behavior by the rules of property”, and Sayre-McCord (2016, 452, 458) argues that it is a concern to do one's share in a mutually advantageous enterprise. I will argue against these readings in the next chapter. The error theory reading is criticized since it leaves our moral beliefs about justice defective, which is, if not necessarily wrong for this very reason, at least unappealing to some commentators.⁹³ Finally, if one would like to adopt the second or third reading by restricting the force of the Circle Argument, one would face the critique that Sayre-McCord (2016, 452) raises against Cohon's reading:

[T]his stands as an interpretation of Hume only if there is reason to think Hume rejected the circle argument when it came to the artificial virtues. If he did, it was – remarkably – without any comment whatsoever.

I believe, nevertheless, that Cohon's interpretative approach is more plausible than the others, and I will defend it against this charge in the next section based on a piece of textual evidence in Hume's argument which has been largely neglected by commentators. By paying attention

⁹³ For instance, see the commentaries of Cohon (2008b, 184–88).

to this piece, we can capture the essence of the natural-artificial distinction addressed in *Treatise* 3.2.1 in a better way.

4.3 The Circle, the Sophistry, and the State of Nature

By paraphrasing Hume's Artificiality Argument in *Treatise* 3.2.1 and locating some problems in it, I argued in the last section that the root of the controversies over its interpretation lies in the lack of a clear sign for the division between nature and artifice in this argument. It seems that Hume starts his reasoning with a proposition that is universally true for all virtues and duties but concludes it by classifying justice as an artificial virtue. To make this argument valid, there must be at least one point in the whole argument where Hume can reasonably announce that 'hereafter what I say only applies to what is natural'. Hume should be blamed for leaving it to his readers to find this point and to account for it, but the situation is actually not that bad because there *is* one place in this argument where he explicitly distinguishes between what is natural and what is artificial. This comes in *Treatise* 3.2.1.9, after the case of the honest man who pays back his debt due to his sense of moral obligation and before Hume repeats the Circle Argument and starts his query for the 'real and universal motive to justice'. There, Hume comments on the motive of the honest man that:

And this answer, no doubt, is just and sufficient to man in his civiliz'd state, and when train'd up according to a certain discipline and education. But in his rude and more *natural* condition, if you are pleas'd to call such a condition natural, this answer wou'd be rejected as perfectly unintelligible and sophistical. For one in that situation wou'd immediately ask you, *Wherein consists this honesty and justice, which you find in restoring a loan, and abstaining from the property of others?* It does not surely lie in the external action. It must, therefore, be plac'd in the motive, from which the external action is deriv'd. This motive can never be a regard to the honesty of the action ... We can never have a regard to the virtue of an action, unless the action be antecedently virtuous. No action can be virtuous, but so far as it proceeds from a virtuous motive. (T 3.2.1.9; 479-80)

This passage, in my view, provides a key to all the puzzles I observed above in this argument, although few scholars pay attention to it. Though this remark comes after the first part of Hume's argument, because there is no substantial difference between this first part, or the Circle Argument, and Hume's briefer argument in the second half of this passage, what can be said of this briefer argument can also apply to the first part of this entire section. In other words, it would not make a difference if Hume had chosen to begin his entire argument with the case of the honest man and the passage cited above and follow it with the longer Circle Argument, the query for the characterising motive of justice, and then the final conclusion. I propose that this

4.3 The Circle, the Sophistry, and the State of Nature

passage indicates that both the Circle Argument and Hume's search for the characterising motive of justice are made under a *false* assumption that justice is a natural virtue. A natural virtue, furthermore, is a virtue which can be approved *in a natural manner* independently of the influence of conventions, and Hume specifies this natural manner with the Motive Principles in this argument. Because the conclusions of the first two parts of Hume's argument are derived from this false assumption, the 'circle' that arises from the clash of them is also restricted under this condition. This explains why Hume concludes his line of reasoning by arguing that:

[T]here is here an evident sophistry and reasoning in a circle. Unless, therefore, we will allow, that nature has establish'd a sophistry, and render'd it necessary and unavoidable, we must allow, that the sense of justice and injustice is not deriv'd from nature, but arises artificially, tho' necessarily, from education, and human conventions. (T 3.2.1.17; 483)

The 'circle' is a piece of *sophistry* or false reasoning⁹⁴ because it starts from a false premise. One avoids this sophistry, consequently, by correcting this premise and acknowledging that 'the sense of justice and injustice' arises artificially. This entire argument, in this reading, is constructed in the form of a *reductio ad absurdum*: by proving that the assumption about the naturalness of justice leads to a sophistry, Hume argues that this assumption is false.

Let me explain in more detail why this reading can be supported by the paragraph above and why it helps with the problems observed in the last section. In *Treatise* 3.2.1.9, Hume distinguishes "man in his civiliz'd state" from man "in his rude and more *natural* condition" by two different reactions people would have to the honest man (who is 'me' – Hume himself and his expected readers – one who lives in a civilized society and has been properly educated) when the latter says that his sense of duty is sufficient to motivate him to repay his debt. Hume distinguishes between, in this sense, two types of conversation partners with the honest man. Those civilized will be *satisfied* with the honest man's answer and will not proceed to ask what motive makes this action just. On the other hand, those who are supposed to live in their natural condition would find the honest man's answer unsatisfactory. To make the honest man's sense of duty intelligible, they appeal to the Circle Argument and require a virtuous non-moral motive, and so begins the search for the characterising motive of justice. If Hume is serious about this contrast between two conditions, this passage can be read as a landmark for the distinction

⁹⁴ In Samuel Johnson's *Dictionary*, 'sophistry' means either 'fallacious ratiocination' or 'logical exercise'. In both the *Treatise* and the *Enquiries*, Hume uses 'sophistry' to refer to a piece of false reasoning (see T 1.1.7.8; 21; EHU 8.2, 12.27, 12.34; 81, 163, 165; EPM 1.1; 169).

4.3 The Circle, the Sophistry, and the State of Nature

between nature and artifice, and it suggests that Hume restricts the validity of the Circle Argument in a supposed ‘state of nature’. This lends support to Cohon’s interpretative approach to which I subscribe at this point.

It is important to observe some characteristics of the state of nature supposed here. Hume does not describe it in detail, so its features can only be inferred from the question of the one who lives in this condition. This man is not literally ‘rude’ in the sense of being ignorant or inexperienced with social life and morality. Indeed, it appears that he has quite a developed ability for reasoning so that he can engage with the Circle Argument which Hume remarks as a piece of “metaphysical subtlety” (T 3.2.1.5; 478), and he must have some experience with sentiment-based moral evaluation to raise this question. Furthermore, he has no difficulty making sense of terms like ‘property’ and ‘loan’. What he regards as “unintelligible”, as can be told from his question which Hume puts in italics, is the *moral quality*, or the “honesty and justice”, of actions to restore a loan or abstain from other’s property, not the *content* of such actions. Interestingly, as this person’s need for a virtuous non-moral motive to account for the merit of justice motivates the search for such a motive, his voice should dominate the entire process of elimination. We see from the latter process that the one who carries it out has full experience of what actions conform with the ‘laws of justice’ and what do not. The secret debt, the debt to a miser, and the rights and obligations the rich hold to the poor, in all these cases, the person who makes the elimination does not doubt what *justice* requires or forbids. He does not argue that since restoring a debt to a “profligate debauchee” only harms the creditor, it is not *just* to restore it; he concludes, instead, that private benevolence cannot characterise just actions. The difference between this person in a supposed ‘state of nature’, on the one hand, and the honest man and other people in their civilized condition on the other, correspondingly, does not lie in rationality or experience but lies in their principles for moral evaluation. It is a state of nature, in other words, of moral sentiments instead of reason or social institutions.

If we compare the state of nature supposed here with Hume’s later account of the origin of justice, it will turn out that this state of nature cannot be historically real. The conventions of property and promises and people’s beliefs of their corresponding rights and obligations, as I shall argue in the next chapter, developed together, and the supposed condition of the person in their ‘natural condition’ is, in a sense, paradoxical as he is civilized in one aspect but uncivilized in another. It is more plausible that this state of nature is a philosophical construction. Later in *Treatise* 3.2.2.14, Hume argues that philosophers can use the idea of a fictional state of nature in their reasoning as long as they do not confound it with the historical

reality. They can make such a “philosophical fiction” by separating the works of different mental faculties which are, by their own nature, inseparable (T 3.2.2.14; 493). The state of nature which Hume appeals to in *Treatise* 3.2.1.9 can then be understood as a fiction that is constructed by *abstracting* some achievement of human artifice from people’s principles of moral evaluation while leaving all other conditions unaltered. This abstraction leads to the conversation partner’s dissatisfaction with the honest man’s moral motive, which further leads to an “unavoidable sophistry”. Since this sophistry arises from a *counterfactual* supposition which is, in a sense, paradoxical, it does not shed doubt on the actual situation of people’s sense of justice and injustice but proves that the latter depends on the piece of artifice which has been abstracted. This may account for the substantial considerations underlying Hume’s entire argument constructed as a *reductio*.

This interpretation helps with all the problems I observed in Hume’s argument in *Treatise* 3.2.1. It suggests that the Motive Principle and the Circle Argument derived from this principle signify features of moral evaluation in a supposed ‘state of nature’. Then, there is no need to worry about the tension between this principle and Hume’s theory of virtue and duty, and Hume has good reason for appealing to this principle here because his final aim is to prove that our sense of justice does not conform with this principle. Second, since the Circle Argument is made under a background supposition about the state of nature, it is no surprise that the conclusion of this argument requires a ‘natural’ or non-artificial motive although Hume does not usually address this point explicitly. The conclusion of the Circle Argument, then, does not differ from the ‘undoubted maxim’, and Hume does not need to explain how he switches from the former to the latter. Third, the same consideration can apply to Hume’s addition of “naturally” to the conclusion of the second part of this argument as the search for the real and universal motive to justice presumes the Circle Argument. It does not matter that the force of this qualification extends to the regard to justice, which, in Hume’s theory, cannot be natural. It is a false claim that we ‘naturally’ have a moral motive to observe the laws of justice, but this is not a problem because this conclusion is derived from a false assumption that justice is a natural virtue or can be approved in a state of nature. Finally, the structure of *reductio*, as I argued above, explains how Hume derives from his first two conclusions that justice is not a natural virtue. The line of reasoning here may be summarized as follows:

- (1) Assumption: Justice is a natural virtue or a virtue that can be praised in a supposed state of nature.

4.3 The Circle, the Sophistry, and the State of Nature

- (2) For all natural virtues, actions can only be virtuous by deriving their merit from a non-artificial virtuous non-moral characterising motive.
- (3) There is no non-artificial virtuous non-moral characterising motive for justice.
- (4) Therefore, the initial assumption is false.

To return to my discussion in the last chapter, this interpretation takes the route which derives C3' from C1'' and C2'. In other words, it inserts a further qualification of 'natural' to both of Hume's first two conclusions by relocating them to a supposed state of nature. It differs from the readings of Garrett and Sayre-McCord because it does not raise a need for an artificial but non-moral virtuous motive to account for the merit of justice. An artificial moral motive, that is, the sense of duty, is sufficient in this aspect, and the task left for Hume's theory of justice is to explain why this is the case. At the same time, though I agree with Cohon on the limit of the force of the Circle Argument, my reading differs from hers insofar as it does not suppose that the artificial virtue of justice needs a characterising motive, and this motive is a sense of duty. The need for a characterising motive follows from the false assumption, not the actual situation of justice. What Hume needs to explain is how the sentiment of moral obligation to justice can work as or produce a motive sufficient for people to observe the laws of justice in a civilized society. It is another issue whether this motive is the only one which can characterise the virtue of justice.

The same interpretation, in my view, can also apply to other arguments in the *Treatise* in which Hume uses the Motive Principle to demonstrate the artificiality of justice. This includes his second argument on promises in *Treatise* 3.2.5.6 and the last paragraph in his third argument in *Treatise* 3.2.6, which I have examined in Section 4.1. In both cases, it seems that Hume derives a dilemma from a *universal* premise about virtues and/or duties but concludes from this dilemma that justice is not a *natural* virtue/duty. It is plausible, based on Hume's distinction between the natural and civilized states in *Treatise* 3.2.1.9, that in both later arguments, Hume overstates the scope of the Motive Principle and the consequent need for a non-moral virtuous motive to make an action virtuous: actually, they only apply to moral evaluation when the influence of convention and education is abstracted from it. It also follows from my reading that the gist of Hume's Artificiality Argument in *Treatise* 3.2.1 is the same as that of his arguments in *Treatise* 3.2.6. What is special to justice, what cannot be observed in common between it and the natural virtues and duties, is the way in which we morally approve of just actions and disapprove of unjust actions. It is especially noteworthy that when Hume first

begins his investigation on the distinction between natural and artificial virtues, he argues that artificial virtues are those which “*produce pleasure and approbation by means of an artifice or contrivance, which arises from the circumstances and necessities of mankind*” (T 3.2.1.1; 477, my italics). Also, when he concludes his argument in this section, he declares that “*the sense of justice and injustice is not deriv’d from nature, but arises artificially*” (T 3.2.1.17; 483, my italics). The bearer of the denomination ‘natural’ and ‘artificial’ is people’s *moral sentiment*, and it is the ‘unnatural’ features of the sentiments we feel about the laws, duty, and virtue which make up Hume’s final *explananda* in his theory of convention.

Before concluding, two further remarks need to be made on my interpretation of Hume’s Artificiality Arguments as a whole. First, there is an apparent difference between his argument in *Treatise* 3.2.1 and those in 3.2.6, namely, in the former, he discusses virtue and duty on the same footing and uses the same argument to distinguish between natural and artificial virtues and duties. In the latter, meanwhile, he argues that all virtues are ‘natural qualities’ while the laws of justice and the corresponding rights and obligations arise from conventions. It might appear that Hume’s view on the relationship between virtue and obligation changes in these sections. This, however, does not need to be the case. What is different here is rather the type of moral obligation which Hume discusses. In *Treatise* 3.2.6, he only examines the sentiment of moral obligation in the case of *justice*, which constitutes the moral relation of property and right and is determined by a set of laws. In *Treatise* 3.2.1, meanwhile, he examines duties that correspond with natural virtues together with the duty of justice. I argued above that he does distinguish virtue and duty at the beginning of this argument when he observes both how we praise actions and how we require them. He applies the same argument to both of them, arguably, because the same feature can be assumed for our moral approbation and sentiment of moral obligation in the supposed state of nature, and both the virtue and the duty of justice do not conform with this feature. This makes the same *reductio* work well for both virtue and duty. Noteworthy, the characteristics of the laws and obligation of justice as highlighted in *Treatise* 3.2.6 are lying in the background when Hume examines the special characteristics of the virtue of justice in *Treatise* 3.2.1. I noted above that some presumed beliefs about what is required by the *laws* of justice play a significant role in Hume’s reasons to eliminate self-love and benevolence from the characterising motive to the virtue of justice. It is presumed that it is clear and uncontroversial what actions are to be specified as just or unjust, and what Hume (in the voice of the man in his state of nature) attempts to find is what motive can constantly produce all such actions as required by the laws of justice so that it can characterise the virtue

of justice. Hume's delineation of what is just and unjust in *Treatise* 3.2.1 does not differ from the universal and inflexible 'laws of nature' that he discussed in *Treatise* 3.2.6, and a proper specification of the virtue of justice, no matter whether it specifies it with the sense of duty as a motive, must accord with the characteristics of such laws. The division between nature and artifice in both sections, consequently, is featured by a contrast between law-independent and law-guided forms of moral evaluations. This applies to both virtue and moral obligation, and Hume does not need to change his mind about their relationship in his arguments.

Second, my interpretation explains how Hume infers the artificiality of justice from the Motive Principle by restricting the latter to a supposed state of nature, and I argued that this state is featured as a 'natural condition' of the moral sentiments. One thing Hume does not explain, nevertheless, is *why* the natural inclination of moral sentiments conforms with the Motive Principle.⁹⁵ I argued in the last chapter that not all virtues in Hume's catalogue derive their moral merit from virtuous motives. It might follow from Hume's presupposition in *Treatise* 3.2.1 that all such virtues should be artificial, but Hume discusses most of them in Book 3, part 3 of the *Treatise*, a part dedicated to natural virtues. It is problematic why, in the supposed state of nature, people were able to approve of benevolence and gratitude, but not good judgement, cleanliness, or the strength of mind. Hume does not mention the Motive Principle in *Treatise* 3.3 or the second *Enquiry*. It seems that his interest in this principle and its theoretical consequences more specifically resides in the Artificiality Arguments in *Treatise* 3.2, but we cannot tell from these arguments his reason for doing so.

It is possible that Hume constructs a 'state of nature' with the Motive Principle out of some polemical reasons. He might have used this argumentative strategy since he intended to object to those who embraced the Motive Principle and thought that justice is a duty characterised by a non-artificial virtuous motive. In the *reductio*, then, he starts with a premise that his opponents would accept and argues that their view leads to a sophistry. If Hume meant his Artificiality Arguments to be addressed against a specific author, in my view, the most plausible candidate is Hutcheson. Noteworthy, Hume's arguments would not be able to persuade philosophers like Pufendorf, Locke, or Clarke, because for them, the moral evaluation of actions can be clearly distinguished from the evaluation of a person's character, and just actions can directly obtain their moral value from their relation to laws or the rational order of nature. They can, consequently, characterise the *virtue* of justice with a sense of duty without falling into a vicious

⁹⁵ I discussed this problem, the difficulties it brings to the interpretation of Hume's conception of 'natural virtue', and probable answers to this question in more details in Qi (forthcoming).

4.4 Conclusion

circle.⁹⁶ Hutcheson would probably agree with Hume's Circle Argument because he endorses the Motive Principle, and because he holds that the moral obligations to abstain from others' property and to perform contracts is grounded on the moral goodness of benevolence.⁹⁷ It is probable, then, that Hume had Hutcheson in mind when writing the second part of his argument in *Treatise* 3.2.1. In a letter Hume wrote to Hutcheson after he had received some comments on his manuscript of Book 3 of the *Treatise* from the latter, Hume repeated the Circle Argument and wrote that Hutcheson did "not assent to this" (L 13; 1:35). We do not know, however, which part of this argument Hutcheson did not subscribe to. It can be told from his *Inquiry* that he agreed with Hume's premises in the Circle Argument (*Inquiry* 2.3.1; 116), and, as suggested by James Moore (1994, 50), Hutcheson would not disagree with the conclusion of the Circle Argument because he did not think that an approval of virtue or sense of duty can account for the merit of an action without referring back to a non-moral motive.⁹⁸ One way to read Hume's words is that he had expected Hutcheson to endorse the Circle Argument or some earlier version of it, and he might have intended *Treatise* 3.2.1 to criticize Hutcheson's view that property and contracts have a natural instead of conventional origin based on premises which Hutcheson and his followers would embrace.⁹⁹

4.4 Conclusion

For many readers of Hume, his theory of moral obligation may first capture their attention with his remarks on it in his Artificiality Arguments, and it may further trigger their curiosity by the paradoxical air in these arguments and the various attempts by scholars to dispel this air and to help Hume escape from the vicious circle. I argue in this chapter, however, that it is not bad for Hume for there to be a paradox in *Treatise* 3.2.1 since the 'circle' is not designed for himself but for those who hold that justice can be praised or required in a supposed 'state of nature' where all moral evaluations of actions observe the Motive Principle. I also argue that all Hume's Artificiality Arguments in Book 3, part 2 of the *Treatise* are dominated by the same problematic, i.e., the way in which 'we', or people who live and were educated in a civilized society, make moral judgements on actions denominated as 'just' or 'unjust' is in tension with

⁹⁶ For instance, Pufendorf argues in his *De Officio* that "Justice is sometimes an attribute of actions, sometimes of persons ... The just man is defined as one who delights in doing just actions or strives after justice or attempts in everything to do what is just ... Justice as an attribute of actions is simply the appropriate fitting of actions to persons." See *De Officio*, 1.2.12-13; 30-31.

⁹⁷ See *Inquiry* 2.7.8; 187-88.

⁹⁸ Also see Norton (2005, 251-54).

⁹⁹ For interpretations which read this Artificiality Argument as made against Hutcheson, also see Moore (1994, 50-54; 2007, 145-46) and Harris (2015, 123-26).

4.4 Conclusion

a supposed natural inclination of the moral sentiments. His different arguments capture different aspects of this tension: our conception of property, its corresponding rights and obligations, the laws to which we appeal, and the reason and motive with which we observe such laws. Moral obligation plays a special role in these arguments because it makes up one of the most significant and salient mark for this tension, and this may explain why the topic of obligation attracts Hume's special attention when it comes to the case of justice. To put it briefly, the tension between the evaluation of justice and other virtues exists because the natural way in which we make sentiment-based moral judgements is not directed by moral laws. Although I shed doubt on Hume's endorsement of the Motive Principle even regarding natural virtues and suggest that he could have designed *Treatise* 3.2.1 to criticize Hutcheson, this doubt does not extend to Hume's other remarks on the natural inclination of the moral sentiments: they are sentiments about persons rather than actions, they are sensitive to particular characters and circumstances of the people principally in concern, and the value they ascribe to persons allow for degrees. The question which confronts Hume here is how to remould notions like law, right, and obligation into non-fundamental moral concepts and integrate them into his theory of virtue and moral sentiments, and Hume responds to this question through his account of the origin of the rules of justice and the sentiment of moral obligation which people feel when observing these rules. This will be the topic of my next chapter.

Chapter 5

Sentiment and Convention: Hume's Method for Artificial Morality

Hume concludes his first Artificiality Argument in *Treatise* 3.2.1 with the claim that “the sense of justice and injustice is not deriv'd from nature, but arises artificially ... from education, and human conventions” (T 3.2.1.17; 483). I argued in the previous chapter that “*the sense of justice and injustice*” here refers to the moral sentiments people feel about justice, including the moral approbation of the virtue of justice and the sentiment of moral obligation to justice or the moral disapprobation of injustice. All of Hume's Artificiality Arguments reveal some aspects in which such sentiments and moral judgements based on such sentiments diverge from the ‘natural’ characteristics of sentiment-based moral evaluation. These ‘non-natural’ or ‘artificial’ features include the preciseness of sentiment-based judgements on the right and obligation of justice, the inflexibility of the laws of justice, and the virtue of justice's exemption from the Motive Principle or the case that just actions and character traits do not derive their moral value from a non-moral motive. As in the *Treatise*, Hume aims to explain the causal principles underlying our sentiment-based moral evaluation, these features require further explanation beyond what can be offered by his basic account of the generation of moral sentiments. This is the task left for his account of “education, and human conventions” which aims at explaining how “the sense of justice and injustice ... arises artificially”.

It is to be questioned, nevertheless, to what extent the conventional origin of justice is important to the explanation of the artificiality of the *sense* of justice. After Hume has finished his first Artificiality Argument and moved on to his positive theory of the origin of justice, he formulates the structure of the latter theory with two questions:

We now proceed to examine two questions, viz. concerning the manner, in which the rules of justice are establish'd by the artifice of men; and concerning the reasons, which determine us to attribute to the observance or neglect of these rules a moral beauty and deformity. These questions will appear afterwards to be distinct. (T 3.2.2.1; 482)

In the interpretations of many scholars, this suggests that Hume takes the conventional origin of justice and the sentiment-based moral evaluation of justice as two 'distinct' topics.¹⁰⁰ They are connected merely extrinsically: conventions give rise to a set of rules and probably also a type of character trait, which we name as 'justice', specified in relation to such rules, but this origin does not influence how people morally evaluate actions and characters of justice. No change would happen to such moral evaluations if the rules of justice had been established in another way as long as people consider justice useful to society. Then, it appears that though Hume's theory of convention can explain how *justice* arises artificially, it cannot explain how *the sense of justice* arises artificially. If there is any human artifice in play in Hume's answer to his second question, it is the education of politicians and parents, "a new *artifice*" which comes onto the stage after the first artifice or convention has done its work (T 3.2.6.11; 533). Hume's theory of convention, though celebrated by many latter-day authors as an important theoretical contribution, then, can be separated from his theory of the *moral* value of justice.

I argue for a different interpretation in this and the next chapter: the conventional origin of justice has a more substantial influence on the moral evaluation people make of justice. The sense of justice arises artificially, then, not only because of its dependence on public and private education but also because the rules of justice are established by conventions. As a special form of social coordination, convention makes its members acquire a specific habit by which single actions related to property or promises are perceived and evaluated as instances of a "general scheme or system of action" (T 3.3.1.12; 579) which is necessary for the interest of the society and every individual. This habit is at work not only in earlier stages of the history of justice but also in larger and civilized societies like those of Hume and his expected readers. Defending this reading is a complex project which I will carry out in three steps. In the first section of this chapter, I argue that such a substantial connection between the two parts of Hume's account of the origin of justice is possible based on his theory of moral sentiments, and that it can be supported by textual evidence. I then dive into the first part of Hume's account, namely, the part on the conventional origin of the rules of justice, to argue that this part can support such a substantial connection. This will take up the remaining sections of this chapter. Finally, in the next chapter, I discuss how the conventional origin of the rules of justice explains the origin of ideas like 'property, right, and obligation', all of which, according to Hume, are only intelligible

¹⁰⁰ For instance, see the readings of Wand (Wand 1956, 160–62), Gauthier (1979, 16–18), Schneewind (1998, 368–72), Don Garrett (2007, 264–72), Cohon (2008b, 171–79), Salter (2015, 34–37), and Chamberlain (2017, 125–30).

after the convention of justice has been established, and how convention-based cognitive and evaluative habits influence the formation of moral sentiments about justice and ideas derived from such sentiments in the second part of Hume's theory of the origin of justice. Since among the three 'fundamental laws of nature', Hume only gives a full account of the origin of the first set of laws, namely, that on the stability of external possessions, I only focus on this part in this chapter, and I follow Hume in using 'justice' to refer to actions, rules, or dispositions to respect others' *property* in corresponding discussions.

5.1 Sentiment and History: Hume's Two Questions about Justice

As I observed above, one crucial question confronted by the interpretation of Hume's theory of justice concerns the relationship between the two parts of his account of the origin of justice. These parts respectively answer his two questions, i.e., how the rules of justice first originate and with what reasons we attach moral value to just or unjust actions. I argue in this section that these two questions, though having 'distinct' concerns, are not 'distinct' in the sense that the conventional origin of justice has no substantial influence on people's moral evaluation of the relevant actions and traits of character. I start with sketching the outline of the two parts of Hume's account of the origin of justice in *Treatise* 3.2.2 and introducing various readings of what the first part of this account provides the second one with. I then appeal to two pieces of textual evidence to argue that one most important connection between these two parts is that the conventional origin of the rules of justice gives rise to an 'artificial manner' in which people recognise the interest served by each single action of justice, namely, such actions are regarded as beneficial to the society and each individual not simply by themselves, but as instances of a 'general scheme of actions' regulated by a set of behavioural rules.

For a brief overview of Hume's answer to his first question, he starts his explanation of the conventional origin of justice by examining human beings' conditions without the rules of property. Their natural infirmity to preserve themselves throughout all necessities of life makes society a necessity for them (T 3.2.2.2-3; 485). At the same time, their naturally selfish and partial passions, together with the scarcity of external goods and the instability of their possession of such goods, induce conflicts and threaten the stability of social cooperation.¹⁰¹ To avoid such conflicts, people "enter into" a convention to "bestow stability on the possession of those external goods, and to leave every one in the peaceable enjoyment of what he may

¹⁰¹ Hume sometimes uses the present tense when discussing the history of justice and sometimes uses past tenses. I will keep using the present tense in my paraphrase in this thesis.

5.1 Sentiment and History: Hume's Two Questions about Justice

acquire by his fortune and industry” (T 3.2.2.9; 489) because they find this policy conducive to the maintenance of the society. The rules of property, then, are first invented with self-interested motives. Insofar as such rules are products of human beings’ “judgements and understanding” (T 3.2.2.9; 489) and serve as a remedy for problems which arise from the natural inclination of human passions, these rules are ‘*artificial*’, though, as Hume clarifies, they are also ‘natural’ insofar as the rational faculty is also a part of human nature and as this invention is “absolutely necessary” to humankind and common to this species (T 3.2.1.19; 484).

After finishing this part, Hume turns to the next topic by remarking that:

The *natural* obligation to justice, viz. interest, has been fully explain'd; but as to the *moral* obligation, or the sentiment of right and wrong, 'twill first be requisite to examine the natural virtues, before we can give a full and satisfactory account of it. (T 3.2.2.23; 498)

The “natural” here means ‘non-moral’, and, although it is controversial what Hume means by ‘natural obligation’ in general, the “interest” here clearly refers back to the self-interest which first motivates the invention of the rules of property. As this remark divides Hume’s responses to his two initial questions, these two parts can also be summarized as respectively focusing on the ‘natural’ and ‘moral’ obligations to justice. Regarding the latter, Hume briefly delineates how moral sentiments arise from sympathy when we are displeased by others’ unjust actions, and how we extend these sentiments to our own actions by following certain “*general rules*” (T 3.2.2.24; 499). While such sentiments are usually too weak to “control our passions”, they can be augmented by a second artifice, i.e., education. As politicians find people’s observance of property rules conducive to their governance, and parents find the character of justice useful to their children, they magnify people’s esteem for justice and abhorrence of injustice “beyond their original bounds” (T 3.2.2.25; 500). The sentiment of moral obligation and other passions related to this sentiment, like the desire for a good reputation which is attached to just conduct by education, then, become forceful motives for well-educated people, and this brings Hume’s entire account of the origin of justice to its close.

Hume does not clarify how these two parts of his story relate to each other. This question has generated scholarly controversies especially since two restrictive considerations make this topic more complicated. First, as I noted in Chapters 3 and 4, many scholars regard Hume as fully committed to the Motive Principle which claims that all virtuous and obligatory actions derive their moral values from their motives. The moral obligation to justice, following this principle, should be an obligation to act with a specific *motive* which can characterise the virtue

5.1 Sentiment and History: Hume's Two Questions about Justice

of justice, but Hume, in the second part of his account of the origin of justice, does not clarify what this motive is. Second, it appears to many scholars that Hume locates his accounts of the natural and moral obligations to justice in two different historical settings. He argues that the regard to self-interest is a sufficient motive for people to act justly “on the first formation of society”. When “society has become numerous”, however, people usually “lose sight of that interest”, and it is because of the wane of the motivational force of the natural obligation that the moral obligation to justice becomes indispensable for maintaining the practice of property in large and civilized societies (T 3.2.2.24; 499). If the object of the moral obligation to justice is a motive, then, it appears unlikely that this motive can be the natural obligation to justice. Based on one or two of these reasons, scholars have suggested the following ways to interpret the connection between the two parts of Hume's account of the origin of justice:

- (a) The conventional origin of justice explains how people find just actions conducive to their self-interest. This ‘enlightened’ self-interest *is* the natural obligation to justice, and it is also the object of the sentiment of moral obligation to justice.
- (b) The conventional origin of justice produces an artificial and non-moral motive to observe the rules of justice which differs from the desire for self-interest, though Hume does not clarify what it is. This motive is the characterising motive for justice and the object of the moral obligation to justice.
- (c) The conventional origin of justice provides objects or contents for a non-moral and uninterested motive which can be felt independently of particular conventional rules, though Hume does not clarify what this motive is. This motive is the characterising motive for justice and the object of the moral obligation to justice.

An instance of reading (a) can be found in the interpretation of Gauthier (1992, 409–25). While ascribing this view to Hume, however, Gauthier argues that Hume fails to provide a stable basis for the virtue of justice because enlightened self-interest is not a reliable motive for justice and that Hume was aware of this difficulty, as can be evinced by his discussion of the case of a “sensible knave” in the second *Enquiry*. One who adopts a knavery policy “conducts himself with most wisdom” and “takes advantage of breaching the laws of justice whenever such actions do not cause considerable harm on the social order” (EPM 9.22; 282–83). Since especially in large societies, not all unjust actions are harmful to public or private interest, one who is good at adopting the most effective means to self-interested ends would choose to act unjustly whenever it is advantageous. Other considerations which shed doubt on

reading (a) are that self-interest is not, by itself, a virtuous or morally obligatory motive, and that this reading makes the virtue of justice undistinguishable from the virtue of prudence.¹⁰²

Don Garrett's (2014, 267) interpretation exemplifies reading (b). According to him, in the development of justice, people first find observing the rules of justice being the best means to pursue one's self-interest. Then, they feel "a new desire to the perceived means", and this desire gradually obtains independency from the original interested desire and becomes a new motive, that is, a "desire and standing disposition to govern or *regulate* one's behaviour by the rules of property", which makes up the characterising motive for the virtue of justice. Garrett, however, leaves three crucial issues unclarified. First, he does not explain how, and in what sense, one's desire for a means can obtain independence from the desire for the original end. Second, since in this model, just actions obtain their value only from the desire for self-interest, it seems that one's desire for such actions is still self-interested, but Garrett holds that self-interested motives cannot be virtuous.¹⁰³ Third, when specifying the virtue of justice, Garrett does not tell apart one's desire to regulate one's actions by the rules of justice and one's 'standing disposition' to do so.¹⁰⁴ As I argued in Chapter 3, in Hume's theory, one's virtuous character traits do not always need to be characterised by a desire. One can have a disposition to regulate one's behaviour by the rules of justice without taking such regulation as the *object* of one's desire. Ambiguities in these aspects, in my view, shed doubt on Garrett's story of the generation of the new, 'artificial' virtuous motive to justice.

Some other scholars who hold that the virtue of justice is characterised by an artificial non-moral virtuous motive embrace reading (c): this motive does not, strictly speaking, arise from conventions but conditions the establishment of conventions. Darwall (1995, 292, 297–98, 308–9), for instance, suggests that the characterising motive for the virtue of justice is a resolution to regulate one's conduct by mutually advantageous rules if others do, and Sayre-McCord (2016, 452, 458) argues that this motive is a concern to do one's share in a mutually advantageous enterprise if others do. Such motives rely on specific conventions, like those about property, to be actually exerted, and they are in this sense 'artificial'. Such motives, nevertheless, are 'natural' in another sense as they, taken as psychological principles, precede

¹⁰² For instance, see Cohon's (2008b, 184–88) criticism of Gauthier's reading.

¹⁰³ In his earlier paper, Garrett (2007, 274) argues that the desire to adhere uniformly to proper rules arises from the belief that this policy is better than all other courses of conduct, and in this sense, it is "caused directly by prospects of pleasure and pain." It then seems that Garrett agrees that this is a self-interested desire.

¹⁰⁴ In his earlier paper, Garrett (2007, 276) formulates the characterising virtuous motive of justice as "a disposition grounded in a desire to regulate one's actions by the rules of justice", which elucidates the relationship between 'desire' and 'disposition' more clearly.

the establishment of the conventions under concern causally. It is questionable whether such motives can find their place in Hume's theory of human nature¹⁰⁵, and if they can, it is still questionable why Hume does not take them into consideration when searching for a non-moral characterising motive for justice in the second part of his argument in *Treatise* 3.2.1. Correspondingly, I doubt whether reading (c) is a promising approach to take.

As I discussed in Chapters 3 and 4, not all interpreters agree that the moral evaluation of just actions conforms with the Motive Principle. Just and unjust actions, in their readings, can be approved or disapproved without referring to their motives, and there is thus no need to take any of the approaches mentioned above to fill in a gap in Hume's theory. According to such scholars, like Mackie (1980, 80), Cohon (2008b, 171), and Harris (2010a, 38–43), what takes place in Hume's account of the origin of justice is that:

(d) After the rules of justice are established by conventions, people morally approve of just actions and disapprove of unjust ones without regard to the motives or characters which produce them.

I take this interpretative approach more tenable. The question which confronts their views (and also mine) is how such an action-oriented type of moral evaluation can be accommodated into Hume's theory of moral sentiments according to which the proper objects of moral evaluation are characters instead of merely external actions. Hume does not explain explicitly how such a reconciliation is possible, and none of these scholars tackles this problem with details in their interpretations. My aim in this and the next chapters is to argue that this gap can be bridged, and the key to resolving this difficulty lies exactly in the first part of Hume's story of the origin of justice. The conventional origin of justice, as I will argue, not only provides 'us', as latter-day moral spectators and agents, with a set of rules to which we refer in moral judgements and motives to observe such rules. It also produces an artificial cognitive and evaluative pattern or habit which influences how we morally evaluate just or unjust actions. By adding this causal connection, I reformulate this more specified version of reading (d) as:

(e) The conventional origin of justice gives rise to a type of habit which makes people regard just actions as morally obligatory without regard to the motives or characters which produce them.

¹⁰⁵ Darwall (1995, 314–18) acknowledges that the motive he suggests, which arises from the 'rule obligation' which he takes as a distinctive type of obligation apart from the natural and moral ones, contradicts Hume's theory of action as it takes rule-regulated acts, instead of some desirable states of affair, as its object.

5.1 Sentiment and History: Hume's Two Questions about Justice

Though a full picture of my interpretation can only become clear afterwards, I would like to first defend its plausibility based on characteristics of Hume's theory of moral sentiments. My reading differs from the first three readings because all of them presume a framework about how the moral sentiments work and search for a factor, like an artificial virtuous motive, to fill up a vacancy in this framework when it is applied to the case of justice, while my reading holds that the conventional origin of justice, to a certain extent, shapes this framework by itself. In this sense, not only the rules and virtue of justice have a historical dimension, but also people's moral sentiments and sentiment-based ideas about them have such a dimension. This is possible because, for Hume, moral sentiments do not arise from an original mental determination like a Hutchesonian moral sense, but are products of more basic psychological mechanisms, that is, sympathy and mental associations along which one's sympathetic feelings are directed and attached to the objects of moral evaluation which, as I noted in Chapter 3, are usually people's mental qualities or traits of character. This psychological process is subjected to the influence of a series of factors. To feel a sentiment of approbation, for instance, an evaluator considers a mental quality's typical useful or agreeable consequences and receive a pleasant sensation from the sympathy with those people who typically benefit from such consequences. This process is guided by causal relations between mental qualities, actions, and their consequences, and the formation of sympathetic pleasure requires the evaluator's recognition of some *non-moral good* that the typical beneficiaries enjoy, as well as awareness of some '*natural relation*' (namely, similarity, contiguity, or causation) between themselves and the beneficiaries with whom they sympathize (T 2.1.11.4-6; 317-18). Moreover, as I discussed in Chapter 3, the formation of the sentiment of moral obligation goes through a more complicated process which may involve an agent's failure to meet some general expectations about virtues or the association between a disapproved vice and its remedy. Custom and habit have a significant role to play in all these nodes: sympathy, causal relations, and the specification of the object of moral obligation guided by general expectations. Although the psychological process which gives rise to an evaluator's token moral sentiment is often transient, the *habits* which influence this process can be moulded by both the psychological development of an evaluator and by the practices and institutions in which evaluators live. As such habits have a history, both on a personal and a collective level, moral sentiments are susceptible of causal explanation in a longer expansion. If the convention of justice can shape such habits, then, the first part of Hume's theory of the origin of justice may lie on the extensive line of the second part, and these two parts may work together to make up one complete causal chain which leads to moral judgements on justice in modern societies.

5.1 Sentiment and History: Hume's Two Questions about Justice

Not only is such a connection between the two parts of Hume's theory of justice possible based on his theory of moral sentiments, but it can also be supported by two crucial passages in which he relies on the conventional origin of justice to explain how people *morally* evaluate just or unjust actions. The first passage is his third argument in *Treatise* 3.2.6 where he observes that the laws of justice are universal and inflexible while the natural inclination of the moral sentiments is sensitive to particular circumstances. Hume accounts for the former by how the rules of justice were first invented:

But 'tis easy to observe, that this wou'd produce an infinite confusion in human society, and that the avidity and partiality of men wou'd quickly bring disorder into the world, if not restrain'd by some general and inflexible principles. 'Twas, therefore, with a view to this inconvenience, that men have establish'd these principles, and have agreed to restrain themselves by general rules, which are unchangeable by spite and favour, and by particular views of private or public interest. (T 3.2.6.9; 532)

Also, to recall, Hume maintains that the solution to the problem of the intelligibility of the moral qualities of property and justice lies in the conventional origin of the laws of justice (T 3.2.6.10; 533), though he does not clarify how this solution works. Another central text is in *Treatise* 3.3.1. When Hume is explaining the origin of moral sentiments, he argues that:

The *only* difference between the natural virtues and justice lies in this, that the good, which results from the former, arises from every single act, and is the object of some natural passion: Whereas a single act of justice, consider'd in itself, may often be contrary to the public good; and 'tis only the concurrence of mankind, in a general scheme or system of action, which is advantageous ... 'twas with a view to this advantage, that men, by their voluntary conventions, establish'd it. (T 3.3.1.12; 579, my italics)

This passage is noteworthy for two reasons. First, it claims that the *only* approach to distinguish natural and artificial virtues is to examine the features of the good, or the useful or agreeable consequences, which a virtue produces. This diverges from the distinction between natural and artificial virtues suggested in *Treatise* 3.2.1, as the definition of natural virtue in *Treatise* 3.3.1 does not involve the Motive Principle.¹⁰⁶ Natural virtues are those qualities the consequences of which arise "from every single act"; while some commentators read this as suggesting that naturally virtuous actions can produce good results at each single instance¹⁰⁷, as Hume opposes

¹⁰⁶ I defended this reading in more details in Qi (forthcoming).

¹⁰⁷ See, for instance, the discussions of McIntyre (2006, 397–98) and Lauren Kopajtic (2015, 214–18) about whether the strength of mind is a natural virtue. Both scholars suggest that a natural virtue should produce good results at each single instance of practice rather than a long-term plan.

this qualification with the case that the advantage of justice arises from “the concurrence of mankind”, it is more plausible to read this “single act” as any action or series of actions which are beneficial *as* actions of individual agents, independently of coordination on a societal level.¹⁰⁸ Hume further delineates natural virtues as virtues the good consequences of which are the objects “of some natural passion”. This qualification differs from the Motive Principle in that it does not imply that the passions under concern should be *virtuous*, or that the virtuous qualities are to be characterised by such passions. It only requires there to be some natural passion which takes the typical consequences of a natural virtue as its objects, which indicates that such consequences are *desirable* by the constitution of human nature. This natural desirability provides ground for the former qualification: when an action or a series of actions aim at a naturally desirable good, it also, for this very reason, tends to be beneficial individually. On the other hand, in many cases, just actions are not beneficial individually, and it is unlikely that they promote or protect an interest which is the object of a natural passion. The artificiality of justice is then featured here by an ‘artificial good’, or a good which can only be recognised in an ‘artificial manner’, that is, as attached to a ‘general scheme or system of action’ which relies on the coordination on a societal level. Arguably, this approach to the distinction between natural and artificial virtues is more consistent with Hume’s virtue theory compared to the one which relies on the Motive Principle based on the doubt I shed on the latter in previous chapters: these two types of virtues differ from each other not by the psychological makeup of the mental qualities which are morally approved but by how people recognise the good consequences of the virtuous actions.¹⁰⁹ Interestingly, Hume only preserves this consequence-oriented approach for distinguishing justice from other social virtues like benevolence in the second *Enquiry* (EPM App.3.2-6; 303-06) and does not mention the Motive Principle in this later work, which further evinces the importance of this approach in his theory.

The second noteworthy point in *Treatise* 3.3.1.12 is that Hume argues that it is exactly with a view of “this advantage”, the advantage which results from “the whole scheme ... of law and justice”, that “men, by their voluntary conventions, establish’d it” (T 3.3.1.12; 579). The same ‘artificial manner’ to recognise the advantage of justice which explains the first establishment of the convention of justice also explains the specific way in which justice is morally approved. Hume employs a similar strategy in *Treatise* 3.2.6.9: we regard the laws of justice as rigid and

¹⁰⁸ On this point, I agree with the reading of Haakonssen (1996, 104).

¹⁰⁹ Abramson (2015) holds a similar view and provides helpful arguments against other interpretations of Hume’s concept of natural virtue.

5.2 History and Reason: Hume's Twofold Account of the Origin of Property

inflexible as these laws are established “with a view to the inconvenience” which arises from the natural avidity and partiality of human passions. These passages support my suggestion that the conventional origin of justice has some significant influence on the manner in which people evaluate just or unjust actions morally, although Hume does not elucidate this influence clearly enough. To see whether such a connection may indeed hold between these two parts, however, we still have to dive into more details in Hume's accounts in both parts. I shall now begin with the first part, on how the rules of justice are first established by conventions.

5.2 History and Reason: Hume's Twofold Account of the Origin of Property

As I argued in Section 5.1, the two parts of Hume's account of the origin of property and justice can be read as a complete causal chain which explains the formation of ‘our’ moral evaluation of just and unjust actions in civilized societies. The first part of this story, namely, how the rules of justice are established by conventions, is crucial to the overarching explanatory project not only because it explains the origin of a set of rules to which people refer in moral judgements or a type of character trait which people regard as virtuous or morally obligatory, but also, and more importantly, because it explains the origin of a specific habit with which people recognise the advantage promoted by just actions. By paying closer attention to Hume's historical account, we can figure out how this habit is first acquired and maintained through different stages in the development of the practice of property. This is the task of the remaining sections of this chapter. In this section, I start with locating Hume's explication of the ‘artificial manner’ in recognizing the interest of justice in the first part of his explanation of the origin of justice. He presents this manner in a model concerning how the convention of property is first invented by an agreement among rational and self-interested individuals. Then, I appeal to Hume's methodological notes on the construction and use of the ‘philosophical fiction’ of the state of nature to argue that this rational agreement should be read as an abstract reconstruction of the ‘natural’ or ‘real’ history of justice; in the latter history, the rules of property evolve under the push of both reason and non-rational faculties like the imagination and natural affections.

As I observed above, in passages like *Treatise* 3.2.6.9 and 3.3.1.12, Hume suggests that it is “with a view to” the advantage that arises from the “general scheme or system” of just actions that people first establish this general scheme by conventions. He emphasizes the same point toward the end of the first part of his account of the origin of justice in *Treatise* 3.2.2, where he notes that:

5.2 History and Reason: Hume's Twofold Account of the Origin of Property

When therefore men have had experience enough to observe, that whatever may be the consequence of any single act of justice, perform'd by a single person, yet the whole system of actions, concurr'd in by the whole society, is infinitely advantageous to the whole, and to every part; it is not long before justice and property take place. Every member of society is sensible of this interest: Every one expresses this sense to his fellows, along with the resolution he has taken of squaring his actions by it, on condition that others will do the same. No more is requisite to induce any one of them to perform an act of justice, who has the first opportunity. This becomes an example to others. And thus justice establishes itself by a kind of convention or agreement; that is, by a sense of interest, suppos'd to be common to all, and where every single act is perform'd in expectation that others are to perform the like. (T 3.2.2.22; 497-98)

The second part of this passage restates Hume's depiction of the first establishment of the convention of justice earlier in the same section:

The convention ... is only a general sense of common interest; which sense all the members of the society express to one another, and which induces them to regulate their conduct by certain rules, I observe, that it will be for my interest to leave another in the possession of his goods, *provided* he will act in the same manner with regard to me. He is sensible of a like interest in the regulation of his conduct. When this common sense of interest is mutually express'd, and is known to both, it produces a suitable resolution and behaviour. And this may properly enough be call'd a convention or agreement betwixt us, tho' without the interposition of a promise ... (T 3.2.2.10; 490)

The invention of the rules of property, as described in these passages, appears to take place in such a manner: every individual member of the society first deliberates on the means to pursue his self-interest and judges that abstaining from others' external possession is conducive to this end, provided that others adopt the same code of behaviour. This causal inference gives rise to a sense of interest which is attached to the observation of the rules of property, rules which now only exist in people's rational deliberation. Then, each person expresses his sense of interest to others and confirms that the same sense is shared across the society. With the belief that others will follow his example, one of them forms a 'resolution' to put his prospective plan about just actions into practice and performs corresponding actions, and others imitate him and bring the "system of actions, concurr'd in by the whole society" into actual existence. In many places, Hume describes this entire series of events as an "agreement" (also see 3.2.2.11, 3.2.3.6; 491, 505), and insofar as people knowingly form resolutions and desires to establish the practice of property, this invention is made by a "*voluntary* convention" (T 3.2.6.11, 3.3.1.12; 533, 579) or agreement. Hume makes it clear that this agreement is not of the nature of a promise, and the voluntariness of this agreement should also have nothing to do with the act of the will which

5.2 History and Reason: Hume's Twofold Account of the Origin of Property

people suppose as attending a promise (T 3.2.5.2, 3.2.5.12; 516, 523). The term 'agreement' rather highlights the state in which all the members of the society agree with each other on their sense of interest and knowingly decide to pursue this interest by a specific type of conduct. I shall follow Hume's usage of this term and refer to the model of the establishment of the rules of property described in these passages as '*agreement-convention*' hereafter.

Though Hume, especially in the *Treatise*, spends much effort in explicating this model of rational agreement, it is controversial to what extent this model reflects his considered view of the development of justice. Haakonssen (1981, 17–18), for instance, argues that Hume's words on this topic in the *Treatise* are not completely self-consistent:

[I]t should be relatively clear that in this passage about the convention concerning justice, we in fact have the indications of two widely different views of the origin of justice. On the one hand what may fairly be called a rationalistic and contractarian view and on the other hand an evolutionary view.

By the "rationalistic and contractarian view", Haakonssen refers to the agreement-convention, and by the "evolutionary view", he refers to a remark Hume makes after introducing the former model:

Nor is the rule concerning the stability of possession the less deriv'd from human convention, that it arises gradually, and acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it. (T 3.2.2.10; 490)

After driving such a distinction, Haakonssen argues that the only view Hume seriously adopts is the evolutionary view, and Hume only "slipped into" the rationalistic view by mistake (1981, 25). The evolutionary view differs from the rationalistic view in two significant aspects. The first is the level of reason involved in the origin of property. The rationalistic model supposes that people are able to grasp property rules as a set of universal laws, which is about to be observed by "all the members of society" (T 3.2.2.9; 489), at the very beginning of the invention of this practice. Every society member is aware that just actions are beneficial to himself only when others act in the same manner, and only after they are convinced about all others' awareness of the same sense of interest can the first person form a "suitable resolution" to carry his intention into action. On the other hand, the evolutionary view does not require the first inventors of property to have such developed rational capacity. Second, the rationalistic model draws a clear-cut distinction between nature and artifice, between a state which lacks property rules and another state where people live under such rules' regulation, and it depicts

5.2 History and Reason: Hume's Twofold Account of the Origin of Property

the convention or agreement as if it is a one-shot event by which people 'convene out' a state of nature where property and justice have no existence. Hume, in many places, writes that the sense of interest people find in observing the rules of property is an interest in exiting a "savage and solitary condition, which is infinitely worse than the worst situation that can possibly be suppos'd in society" since "without justice, society must immediately dissolve" (T 3.2.2.22; 497). The evolutionary view, meanwhile, does not suppose such a distinction.

Haakonssen (1981, 17–26) supports his doubt on Hume's endorsement of the rationalistic model of the origin of justice by pointing out its inconsistency with Hume's other remarks on the same topic. For example, Hume argues that "the very first state and situation" of human beings is "social" because their most natural passions, that is, sexual desire and love of one's offspring, unite them into families (T 3.2.2.4; 486). Some "first rudiments of justice" are already in place in such smallest and most natural social units, and a pre-social state of nature, according to Hume, is merely a "philosophical fiction, which never had, and never cou'd have any reality" (T 3.2.2.14; 493). In families and probably also slightly larger communities, people learn to constrain their natural desires by property rules "upon the least reflection" (T 3.2.2.13; 492); they do so only with a view to their own interest (T 3.3.1.9; 577), and the society benefits from justice only as an unintended consequence (T 3.2.6.6; 528-29). As time goes on, the practice of property expands to a wider scale, and the rules of justice become more rationalized and universal. People's recognition of such universal rules and their contribution to the public interest then arises as a result of the development of justice instead of preceding it as a cause, and there is, strictly speaking, no clear-cut distinction between nature and artifice because the artificial rules of justice emerge and develop 'naturally', that is, under the push of natural passions without much precedent rational and intentional design. The distinction between nature and artifice, according to Haakonssen (1981, 23), "is really a non-distinction" in Hume's considered view.

Similar discrepancies between two accounts of the origin of property in Hume's theory are also spotted by scholars like Gerald Postema (1986, 114–21), John Latsis (2009, 220–32), and Harris (2012, 219).¹¹⁰ I agree with them that Hume indeed presents two stories about how justice first arises from conventions, and in the first part of the section on the origin of justice, he switches between them without reminding his readers about their differences. To see what

¹¹⁰ Harris embraces Haakonssen's view that the model of gradual development more plausibly expresses Hume's considered view, while Postema argues that these two views describe 'two sides of the same reality' respectively in analytical and historical aspects.

role the agreement-convention plays in Hume's theory of justice in general, then, it is of utmost importance to clarify the nature and function of this account in Hume's twofold story, especially in relation to what Haakonssen calls the 'evolutionary view' or the model of gradual development.¹¹¹ In this aspect, I disagree with Haakonssen that Hume constructs the model of agreement-convention because he fails to stick to his own considered position out of mistakes. Instead, textual signs in the *Treatise* suggest that Hume offers such a twofold story intentionally, and his two accounts delineate the same historical process with two different methods. To see this, more attention needs to be paid to Hume's discussion of the 'philosophical fiction' of the state of nature, concerning how such fictions are constructed and what functions they may serve in his theory of justice.

As I argued in the last chapter, though Hume stresses that the fictional idea of the state of nature should be strictly distinguished from historical reality, he does not deny that it is a useful tool for philosophical reasoning, and he himself makes use of this tool. One can construct a scenario of the state of nature, as he argues, by the method of separation:

Human nature being compos'd of two principle parts, which are requisite in all its actions, the affections and understanding; 'tis certain, that the blind motion of the former, without the direction of the latter, incapacitate men for society: And it may be allow'd us to consider separately the effects, that result from the separate operations of these two component parts of the mind. (T 3.2.2.14; 493)

Moral philosophers can separate different mental faculties and observe their work in isolation, just like natural philosophers take a motion as compounded and separate it into parts although this motion is "in itself uncompounded and inseparable" (T 3.2.2.14; 493). When making this piece of methodological note, Hume's focus is on a state of nature which is "describ'd as full of war, violence, and injustice" (T 3.2.2.15; 493) and which stays in contrast to his position that the first state of human beings is social. Hume's methodological note indicates that this state is constructed by abstracting the work of understanding from the affections, which leaves the latter moving "blindly". In the second *Enquiry*, Hume discusses the same scenario of the state of nature and ascribes it to Hobbes and Cicero. People in this condition are supposed to be entirely selfish, ignorant, and incapable of mutual trust (EPM 3.15; 189-90).¹¹² In both the

¹¹¹ Hume does not use 'evolution' to describe the history of justice, and I shall hereafter use 'gradual development' instead of 'evolution' to avoid some unnecessary theoretical implications of the latter term.

¹¹² Hume might have misread or misrepresented Hobbes because people in Hobbes's state of nature are not slaves of blind passions. They are rational enough to anticipate others' potential threats to their life and to listen to the

5.2 History and Reason: Hume's Twofold Account of the Origin of Property

Treatise and the second *Enquiry*, Hume puts this philosophical fiction together with the “poetical fiction” of a “golden age”, in which people are unacquainted with avarice, selfishness, and the scarcity of resources. Both fictional states are counterfactual, and in both states, property would not be invented: it would not arise in the alleged Hobbesian state of nature because people there lack reason and experience of sociable life, and it would not arise in the golden age because people there lack the motive to do so (T 3.2.2.15; 493-94; EPM 3.16; 190). In this way, both counterfactual states work to highlight some necessary conditions for the establishment of the convention of justice.

Though this instance of Hume's use of the state of nature is widely discussed in literature, not many scholars pay attention to other places where Hume uses the method of separation in different manners for different goals. I discussed in the last chapter that Hume resorts to a fictional figure, a person “in his rude and more *natural* condition” (T 3.2.1.9; 479), in his first Artificiality Argument. This person can engage with subtle philosophical reasoning and has knowledge about the institutions of property and promises, but he takes the moral value which people in their “civiliz'd state” ascribe to just actions as unintelligible. I argued that Hume makes up this figure by the same method of separation, though he does not suppose this person in the state of nature as devoid of *rationality* but abstracts from him the type of moral *sentiments* which rely on convention and education. This person's condition is also counterfactual, and Hume uses it to demonstrate the reliance of our ‘sense of justice’ on human artifice.

To return to the agreement-convention, I suggest that it can be read together with a later passage in *Treatise* 3.2.3, where Hume describes the condition of the first inventors of justice:

I first consider men in their *savage and solitary condition*; and suppose, that being sensible of the misery of that state, and foreseeing the advantages that wou'd result from society, they seek each other's company ... I also suppose, that they are *endow'd with such sagacity* as immediately to perceive, that the chief impediment to this project of society and partnership lies in the avidity and selfishness of their natural temper; to remedy which, they enter into a convention for the stability of possession, and mutual restraint and forbearance. (T 3.2.3.3; 502-03, my italics)

This condition diverges from Hume's view about the “very first state and condition” of human beings in which they are born in families and acquire rationality in society. It is clearly another philosophical fiction because Hume supposes people in this condition as solitary but sagacious,

dictate of right reason, which partly enables them to ‘contract out’ the state of nature. For Hobbes's method to construct the state of nature, see especially *De Cive*, 10-12, 21-31. Correspondingly, I refer to this state in Hume's works as an ‘alleged’ instead of a truly Hobbesian state of nature.

living without society but fully aware of the benefit of society and the method to overcome the “chief impediment” to such a project. At the same time, this fiction also differs from the alleged Hobbesian state of nature where human beings are supposed to be devoid of understanding and incapable of cooperation.¹¹³ It also differs from the latter in its function. While the alleged Hobbesian state of nature, like the golden age, are counterfactual scenarios which work to show what conditions are indispensable for the invention of property in reality, the state of nature in *Treatise* 3.2.3.3 is one which people both need to and can exit by ‘entering into’ a convention, and Hume employs it to illustrate how property rules *are* invented. Though not addressed in the text, this scenario of a pre-conventional state of nature fits best with the agreement-convention compared to all other descriptions of human beings’ condition before the invention of property, both concerning the degree of rationality involved in this invention and the comparably clearer distinction between nature and artifice.

After introducing his supposition concerning the inventors of justice, Hume also clarifies the relationship between this supposition and the ‘real’ history of justice:

I am sensible, that this method of proceeding is not altogether natural; but besides that I here only suppose those reflections to be *form'd at once*, which in fact arise *insensibly and by degrees*. (T 3.2.3.3; 503, my italics)

This remark conveys Hume's methodological self-awareness when writing about this pre-conventional state of nature, and it explicitly shows that he has in mind two methods to present the development of the institution of property and that their key difference lies in the way how people's rational reflections on the inconvenience caused by their natural avidity and partiality and the remedy to this problem arise. In the more “natural method of proceeding”, such reflections emerge “insensibly and by degrees”, while in this fictional state of nature, they are described as formed at one stroke. This can be read as another application of the method of separation. It aligns with Hume's study of human nature that in the real world, people's rational reflections on the interest of justice evolve together with their passions, sentiments, and habits. The operations of these faculties are closely intertwined, and the improvement of reason, both for individuals and for entire communities, is usually both conditioned by suitable work of the

¹¹³ Though Forbes (1975, 69–80) and Postema (1986, 92–93) pay attention to this scenario of the state of nature, they read it as the same fiction as the alleged Hobbesian one, to which I disagree. This ‘Humean’ state of nature is indeed more similar to Hobbes's own supposition about the state of nature than the alleged ‘Hobbesian’ state of nature, although they still differ in many aspects. Paul Sagar (2018, 96–99) distinguishes between a historical and an analytical and hypothetical state of nature in Hume's discussion and uses this distinction to reply to Haakonssen, but he does not mention T 3.2.3.3 (503) and addresses the analytical idea with T 3.2.2.14 (493).

other faculties and, in turn, of significant influence on the latter. The progress of these faculties is in fact inseparable. In the “unnatural method of proceeding”, the supposition that the rational reflections on the necessity of justice are “form'd at once” can be read as a result of separating the achievement of reason from the contribution of non-rational mental faculties, which makes it appear that the former arises simply by itself, out of nowhere, and at one stroke. This coheres with Hume's suppositions about the pre-conventional state of nature. As Hume argues, when deprived of the guidance of reason, passions “incapacitate men for society” (T 3.2.2.14; 493), which accords with the supposition that human beings in this condition are “savage and solitary” like in the alleged Hobbesian state of nature. At the same time, Hume does not deprive these people of rationality but puts the work of reason, now isolated from the passions, back to the scene, which explains why people in this condition are rational. The final product is a fictional scene in which solitary rational individuals deliberate about their self-interest and exit a pre-social state of nature by entering a convention or agreement concerning the stability of external possessions.

It follows from this interpretation that the model of agreement-convention should be read as a philosophical reconstruction of what could ‘in fact’ have happened in the history of justice. Importantly, this indicates that this model about the origin of justice is partly *fictional* and partly *real*. By ‘real’ here, I do not mean what can be supported by testimonies like written history or travellers' reports; Hume's historical narrative in the *Treatise* is more of a product of theoretical analyses and conjectures.¹¹⁴ Hume opens up the space for historical narrative with his Artificiality Arguments: property and justice *need* a history because this practice, especially people's moral evaluations about it, does not completely agree with the natural inclinations of human beings. His account of the situation of human beings before the full establishment of the convention of property is based on analyses of human nature and their circumstances, and the plausibility of this account relies on an underlying presupposition about the uniformity of

¹¹⁴ Dugald Stewart regards Hume as a predecessor of the approach of ‘conjectural history’ which was later adopted by a group of Scottish authors. Hume's account of the origin of justice in the *Treatise* may count as a conjectural history according to Stewart's specification that it is a practice which one may appeal to “when we are unable to ascertain how men have actually conducted themselves upon particular occasions” and which one attempts by “considering in what manner they are likely to have proceeded, from the principles of their nature, and the circumstances of their external situation”, though Stewart's point of reference, when making this claim, is Hume's *Natural History of Religion*. I do not use the phrase ‘conjectural history’ here because it is controversial to what extent Hume's approach to the history of justice in the *Treatise* resembles the approaches employed by authors like Kames, Smith, and Stewart. For different views, see the discussions of Simon Evnine (1993, 589–604) and Juan Samuel Santos Castro (2017, 163–70).

human nature throughout history and across different societies and customs.¹¹⁵ When Hume says that people's rational reflections on the interest of justice "*in fact* arise insensibly and by degrees" (T 3.2.3.3; 503, my italics), then, he should be taken as suggesting that it is more probable, based on his studies of human nature, that people acquire such reflections gradually, and that such a process is more probably part of the causal chain which extends from the "first situation of human beings" to their current and civilized condition and which can thereby give a valid explanation for the generation of their sense of justice. This is the sense in which I talk about what is 'real' about Hume's historical narratives, and in this sense, property rules are not 'really' established by a process like the agreement-convention since it follows from Hume's theory of human nature that a state in which people are both "solitary" and "endow'd with sagacity" could hardly exist.

At the same time, my reading of the nature of the agreement-convention also indicates that this fictional account is constructed *based on* reality, or what can be supposed as the reality of the history of justice. This account is 'real' insofar as its suppositions about human nature, about the inclination of their passions and the degree of rationality that the inventors of justice could have, are derived from empirical observation, and also insofar as the *contents* of people's reflections on the interest of justice do not differ in Hume's two accounts of the origin of justice; these accounts diverge only in the manner in which such reflections arise. In this sense, the agreement-convention still represents what could 'really' have happened in history; though being a rational reconstruction, it is not completely ahistorical. It is implausible, then, that this model is intended to serve as a *hypothetical* account in which idealized agents deliberate, bargain, and reach agreement on observing a set of conventional rules, though some scholars, like Gauthier (1979, 7–29), read Hume's theory of property as in alignment with hypothetical contractarianism.¹¹⁶ Only insofar as the agreement-convention represents certain aspects of the 'real' history of justice can this account have a place in Hume's causal explanation of the generation of moral sentiments about justice. Now that the nature of the agreement-convention has been clarified, we can move on to discuss what this model of the origin of property, when

¹¹⁵ For different readings of Hume's view on the uniformity of human nature, see, for instance, the discussions of Forbes (1975, chap. 4) and Evnine (1993, 589–604). I agree with Forbes's view that Hume subscribes to the uniformity of the psychological principles which he establishes by observation and induction, and uniformity on this level is compatible with diverse expressions and consequences of such principles in different societies. Evnine's (1993, 602–4) point that human rationality went through significant development in Hume's account of the origin of justice, in my view, is indeed consistent with Forbes's interpretation as this development can be accounted for by the work of custom rather than a change in the rational faculty itself.

¹¹⁶ For criticisms of Gauthier's reading, see, for instance, the discussions of Castiglione (1994, 110–11), Frederick Whelan (1994, 216–18), Christine Chwaszcza (2013, 111–19), and Sebastian Bender (2022, 110–11).

put back to its proper place in Hume’s conjectural historical account, indicates about the development of the ‘artificial manner’ in which people recognise the interest of justice.

5.3 Reason and Custom: The ‘Real’ History of Justice

I argued in Section 5.1 that Hume’s theory of moral sentiments allows that not only the rules and virtue of justice have a history, but also the way in which people morally evaluate just and unjust actions. The latter can be shaped by the conventional origin of justice insofar as this long historical process can give rise to specific habits with which people recognise just actions and their benefit to private and public interest, and such habits can in turn influence how they make sentiment-based moral judgements about justice. One crucial clue here is Hume’s remark that the only difference between justice and the natural virtues is that just actions are only regarded as useful when considered as instances of a “general scheme of actions” with the “concurrence of mankind” (T 3.3.1.12; 579). In Section 5.2, I traced Hume’s words on this ‘artificial manner’ to recognise the interest of justice in his historical account: he discusses this sense of interest as a first motive with which people enter into an agreement to establish the rules of property. By discussing scholarly commentaries on this account, which I call *agreement-convention*, and reading it together with Hume’s methodological notes on the construction and use of the state of nature in philosophical reasoning, I argued that this account is a fictional reconstruction of the ‘real’ history of property and justice. Instead of being a loyal description of the development of justice, it represents this process by isolating the contribution of reason in it. In this section, I look through this account to see what is *represented* by it, or how this ‘artificial’ recognition of interest first arises and is maintained throughout Hume’s historical narrative in light of his more ‘natural method of proceeding’. I shall argue that: (i) the rules of justice are joint products of people’s passions, imagination, and reason; in this process, which I shall refer to as ‘*custom-convention*’, people gradually obtain general ideas about justice and general rules concerning the causal relation between justice and private and public interest; (ii) in relation to the custom-convention, the agreement-convention works to present the special form of such general rules.

Although Hume declares his awareness of the case that the agreement-convention and its corresponding state of nature are depicted in an ‘unnatural method of proceeding’, and although he claims that the invention of property ‘in fact’ takes place gradually, he does not give many words to the latter ‘real’ and more fundamental process, especially in the first part of *Treatise* 3.2.2. The agreement-convention takes up a major part of his attention there. Accordingly, there are not many clues about what other principles in human nature, apart from the desire for self-

interest and instrumental reasoning, also play indispensable roles in the development of justice. Limited clues in the *Treatise* include Hume’s remarks on the importance of family and the education of parents for people to obtain their first experience of the benefit of society, sociable affections, and rudimentary forms of justice (T 3.2.2.4, 3.2.2.14; 486, 493), as well as his words on the “jealousy of interest” which “justice supposes” (T 3.2.2.16; 494) and which can be explained by people’s natural inclination for comparison. In the second *Enquiry*, he also talks about the ability to make the effects of one’s resentment felt as a necessary condition for one to take part in the practice of justice (EPM 3.18; 190) and argues that justice cannot be invented without ‘mutual regard and forbearance’ among people (EPM 3.20; 191).¹¹⁷ Familial affections, some basic mutual regard among strangers, resentment, and jealousy, on these accounts, are all parts of the psychological mechanisms which condition, motivate, and shape the interactions in which the convention about property gradually emerges.

Besides these passions, another faculty which is necessary for the invention of property is imagination. Hume makes lengthy discussions on the contribution of custom and imagination to the establishment of particular property rules in *Treatise* 3.2.3; their influence makes certain relations between persons and external objects, like present possession and accession, more salient to people, and makes it easier for them to ‘acquiesce in’ taking such relations as the bases for property relations (T 3.2.3.4, 504-05). Though Hume discusses the invention of particular property rules after the part on the ‘general rule, *that possession must be stable*’ (T 3.2.3.3; 502), the latter making up the proper topic for *Treatise* 3.2.2, it is implausible that this order of narrative could be temporal according to the ‘natural’ history of justice since it is unlikely that the general rule about the stability of external possessions could be actually established before having any specific content.¹¹⁸ More probably, custom, imagination, as well as natural passions should have exerted their force *before* people could reflect on the interest of justice understood as a general scheme regulated by universal laws. Moreover, the influence of these factors should *accompany* the development and maintenance of the practice of property over time, since the natural inclinations of such faculties do not change as long as human nature keeps uniform. Instead of reading Hume as providing a *two-stage* account of the first invention of justice, as argued by scholars like Simon Evnine (1993, 604) and Juan Samuel Santos Castro (2017, 164), in which non-rational faculties and rational reflections *take turns* to contribute to

¹¹⁷ For the importance of family, natural passions like resentment, and some natural virtues to the development of the convention of justice, especially see the discussion of Baier (2010, chaps 3, 6, 7).

¹¹⁸ For a similar view, see Postema (1986, 121–22).

the development of justice in different historical stages, I suggest that Hume provides a *two-layer* account, each layer described by his ‘natural’ and ‘non-natural methods’ respectively. The interaction of passion, imagination, as well as reason and custom in conflicts over external possessions and attempts to settle them makes up a more fundamental layer of the development of the rules of justice, and the account of agreement-convention is abstracted from this more fundamental layer and presented as a reconstruction of this process, in which the work of reason is isolated and highlighted.

I shall refer to the process in which the rules of property gradually develop under the push of both non-rational faculties and rational reflections as ‘*custom-convention*’, as a counterpart to the model of agreement-convention in Hume’s rational reconstruction. I use the term ‘custom’ in this denomination to stress that many parts of this process usually take place pre-reflectively. “[W]e call everything CUSTOM,” as Hume suggests, “which proceeds from a past repetition, without any new reasoning or conclusion” (T 1.3.8.10; 102).¹¹⁹ Hume relies on custom to explain the formation and use of abstract ideas (T 1.1.7.7; 20), beliefs based on causal relations (T 1.3.8.9-13; 102-104), judgements of probability based on general rules (T 1.3.13.7-9; 146-48), and the increase of the strength of calm passions when it “has once become a settled principle of action” (T 2.3.4.1; 419), all being psychological processes which “may operate on our mind in such an insensible manner as never to be taken notice of, and may even in some measure be unknown to us” (T 1.3.8.13; 103). Through all these mechanisms, custom interacts with passions, imagination, and reason and explains why people can obtain new cognitive and evaluative patterns of habits “insensibly and by degrees”.

As the agreement-convention is an abstraction of the custom-convention, its proper function in Hume’s account of the history of justice can only be clarified when read in relation to the latter, more fundamental process. Importantly, clarifying this relationship helps to avoid some misreading about the inventors of property in their ‘real’ historical situation. Instead of being atomic, rational individuals, these people are passionate beings within various types of social relations. They learn to distinguish *mine* and *thine* from repeated experiences of conflicts over external possessions, and in this way, they gradually become individual or joint proprietors

¹¹⁹ In the first *Enquiry*, Hume defines custom as ‘a principle of human nature’ which is ‘well known by its effects’: ‘wherever the repetition of any particular act or operation produces a propensity to renew the same act or operation, without being impelled by any reasoning or process of the understanding, we always say, that this propensity is the effect of *Custom*’ (EHU 5.5; 43). In the following discussions, I shall use ‘custom’ for the underlying mental principle, ‘habit’ for its products, and ‘customary’ for the quality of mental transitions which arise from custom, though Hume does not stick to such a distinction.

of certain objects. Also, what is described in the agreement-convention as people’s *prospective* instrumental reasoning should result from their *retrospective* reflections on past experience, and when people make such reasoning, they should already have obtained abstract ideas about what counts as one’s property and general rules about the probable consequences of actions which violate one’s property. It is a piece of misreading, correspondingly, to suppose that by the time the agreement-convention is first established, every *single case* of just or unjust action must produce some benefit or harm to private or public interest; the causal relation between justice and interest should already have been grasped in a generalized manner at this point. Neither does the agreement-convention indicate that at some earlier stage in history, single cases of injustice could indeed make the society dissolute and make people, or at least the rule-breaker, fall back to a ‘savage and solitary’ state of nature. I argued above that this state of nature is only a fictional scenario with which Hume constructs the agreement-convention. The sense of interest which people feel about observing the rules of property should rather also be derived from their generalized thoughts based on experiences of conflicts over property; it is such experiences, instead of the fear of the state of nature, which drive people to respect others’ property when the rules of justice are first established.

If the agreement-convention does not reflect the ‘real’ history of justice in all these aspects, then, why does Hume put so strong emphases on this account in the *Treatise*? In my view, this is because by isolating the work of reason in the gradual development of property rules, Hume summarizes the special *form* of the ‘*general rules*’, or customary associations between causes and effects, that people follow in grasping the interest of justice, although such general rules in fact arise from custom and more usually exert their influence in pre-reflective cognitive and evaluative activities.¹²⁰ In this sense, I take the agreement-convention as presenting a *formula for generalization* which people adopt when perceiving the interest of justice. This account is indispensable to Hume’s explanatory project in the *Treatise* as it distinguishes the general rules about the interest of justice from other rules of causation or probability which people follow when relating actions to their consequences. To recall, in the agreement-convention, everyone is aware that his observation of property rules is only beneficial to his interest *provided that* other people adopt the same code of behaviour. This structure of mutual conditioning may arise from people’s experiences about conflicts over external possessions and attempts to settle such

¹²⁰ For Hume’s technical use of the phrase ‘general rules’, which he usually italicizes in the *Treatise*, in the sense of instinctive causal inference, and the crucial role ‘general rules’ play in the formation of moral sentiments, see the discussions of Gill (2006, 219–24; 2014, 27–28), and Chamberlain (2022, 83–90; forthcoming).

conflicts. From the perspective of the custom-convention, such experience gives rise to habits which make people consider single actions to abstain or violate others’ external possessions in relation to similar actions performed by *other* community members when they think about the causal relation between such actions and the interest served by them. Since especially when the practice of property is first invented, people are continually pushed by their interest to think about what types of others’ external possessions they should abstain from, this structure of the mutual reference of similar actions shapes the way in which they obtain their very first general ideas of ‘property’ and ‘justice’. When they form customary general rules about the interest of justice, likewise, they do not consider the causal factors in such relations merely as actions of individual agents but consider such actions, insofar as they are ‘just’ or ‘unjust’ actions related to someone’s ‘property’, as *instances* of a ‘general scheme’ constituted by similar actions. This form of general rules differs from those concerning the useful or agreeable tendencies of actions like benevolent or prudent ones since the latter can be related to their ends without relying on the coordination of others on a societal level. This, to recall, is the “only” difference between justice and natural virtues which Hume elucidates in the passage on the consequence-oriented distinction between natural and artificial virtues.¹²¹ Hume resorts to the agreement-convention to explain the artificiality of justice, then, not because this account describes a series of events which take place in the ‘real’ history of justice, but because it is a reconstruction of this historical process which characterises the special formula for generalization, or the special form of the general rules, with which people relate just or unjust actions to their interest.

At this point, we can finally see how Hume’s account of the conventional origin of justice may provide his causal explanation of the artificial features of the sense of justice of what the latter needs. By clarifying this relationship, meanwhile, we are still only halfway through our journey because two questions remain to be answered. First, I have not discussed in this chapter how people’s moral sentiments about justice relate to their recognition of the interest of justice. Since the moral obligation to justice is grounded in moral sentiments, and since such sentiments, though arising from people’s recognition of some non-moral good, cannot be reduced to the latter, an account of the manner in which people recognise the interest of justice cannot suffice to explain characteristics of the *moral* obligation to justice. Second, I have not discussed how

¹²¹ Consequently, I disagree with Taylor (1998, 23–24), Gill (2014, 27–28), and Chamberlain (2017, 125–26, 129–34) that exactly the same influence of general rules which explains the moral approbation of ‘virtue in rags’ (T 3.3.1.20; 584–85) also explains the moral approbation of single just actions when they do not contribute to public or private interest by themselves. In my reading, though Hume resorts to the work of general rules to explain both psychological processes, the forms of the general rules, or the pattern in which ideas of actions and their consequences are generalized, are different in these cases.

Hume's account of the origin of justice, which addresses a process which can only be supposed as taking place in the remote past of human societies, can explain the characteristics of moral judgements about justice of 'us' 'now', or in large and civilized societies, where the forms of social coordination and people's situations have changed vastly compared to the small societies in which the rules of justice first emerged. These two topics, in Hume's account of the origin of justice, are closely related to each other, and I shall turn to them in the next chapter.

5.4 Conclusion

I discussed in this chapter how Hume's historical account of the conventional origin of justice fits into his overarching project in the *Treatise* which intends to explain the 'artificial' features of our sentiment-based moral evaluation of justice. I argued that Hume's two questions about justice, namely, how its rules are first established by human artifice and with what reason we regard the observation of such rules as morally good, have a more substantial connection than how they are read by many scholars. To defend this interpretation, I looked into both sides and identified their connection by shedding light on each side's relationship to custom and habit: the formation of moral sentiments and sentiment-based ideas can be influenced by customary mental associations, and Hume's central account of the conventional origin of justice, namely, the account in which property rules are established by an agreement among rational and self-interested individuals, can be read as an abstract formula for customary mental associations or general rules about the interest of justice. Now as both parts of Hume's theory of the origin of justice take a step toward each other, it becomes clearer how they can make up one complete causal chain which connects the first invention of property rules to the formation of relevant moral sentiments.

Before moving on to the next phase of Hume's account of the origin of our 'artificial' sense of justice, we can already observe from the discussions above what this 'artificiality' indicates. Hume clarifies at the very beginning of his investigation of the principles of moral sentiments that by 'artificial', he means things which owe their existence to "the designs, the projects, and views of men" (T 3.1.2.9; 474); he claims later that justice is an artificial remedy for problems which arise from human nature as it is invented by "judgement and understanding" (T 3.2.2.9; 489), and he notes in the second *Enquiry* that what is artificial is what supposes "reason, forethought, design, and a social union and confederacy among men" (EPM App.3.9n; 307). In all these texts, Hume annotates the artificiality of artificial virtues by humankind's rationality and inventiveness, with which they depart from their natural instincts and inclinations, reflect

5.4 Conclusion

on their situations, and establish social practices to serve their self-interest. When Haakonssen (1981, 20–26) sheds doubt on the consistency between Hume’s account of the agreement-convention and remark on the gradual development of justice, he suggests that the distinction between nature and artifice in Hume’s theory is a “non-distinction” because strictly speaking, justice and property emerge from human beings’ natural inclinations and social interactions without forethought and design. They arise spontaneously and promote the public interest as an unintended consequence, which, according to Haakonssen, is an important step that Hume takes away from previous authors who regard social and political norms as a human creation. The coexistence of two accounts of the origin of justice, i.e., the custom-convention and agreement-convention, in my view, shows that Hume’s considered view on the implication of the ‘artificiality’ of justice is between these two extremes. Property and justice are not products of people’s deliberate design as they first take shape through pre-reflective activities guided by natural passions, imagination, and custom. As time goes on, meanwhile, each person, according to Hume, obtains rational reflections on the interest of justice and learns to construe just actions as instances of a ‘general scheme of actions’ regulated by universal laws, and such reflections, in turn, gradually influence people’s pre-reflective activities by customary mental associations. In this sense, property and justice are joint products of non-rational faculties and reason, of a more *fundamental* process in which each person’s natural passions and imagination clash and reconcile with each other and a more *derivative* process in which they reflect on the former and agree to continue to abide by norms which emerge from the former. By saying that the rules of justice are artificial in the following parts of this study, I use ‘artificial’ in this modified sense, not only addressing the work of judgement and understanding but also addressing their work *as* resulting from and reflecting the dynamics of passions and imagination in social interactions. As I will discuss in the next chapter, our idea of the moral obligation to justice is also the joint products of this artifice and the natural inclination of our moral sentiments.

Chapter 6

From Interest to Morality: Convention and Obligation

Hume suggests in the *Treatise* that the discrepancies between the way we morally evaluate just actions, on the one hand, and the natural inclination of our moral sentiments on the other can be explained by how “the sense of justice” arises “from education, and human conventions” (T 3.2.1.17; 483). He then separates this account into two parts, respectively about how the rules of justice were first established by convention and how we give moral value to the observance and neglect of such rules (T 3.2.2.1; 484). I argued in the last chapter that, unlike how this is widely interpreted by scholars, the conventional origin of justice has a substantial influence on the moral evaluation of justice because moral sentiments and sentiment-based judgements are formed under the influence of custom and habits, and the formation of such habits allows for causal explanation with a historical dimension. Then, I argued that Hume delineates the history of justice with a twofold account in which he describes the same conjectural process from two different aspects, based on two sets of methodological suppositions about human nature before the invention of justice. One of them, which I called the model of *custom-convention*, describes the dynamics of passions, imagination, custom, and rational reflections in social interactions and shows the more fundamental aspect of Hume’s historical account. The other, which I called the model of *agreement-convention*, is a derivative and selective representation of the former in which Hume analyses the structure of the convention of justice and underscores the special form of convention-dependent *general rules* by which convention members relate just or unjust actions to the interest which these actions, insofar as they are *just* or *unjust* actions, promote or endanger.

This chapter offers the second part of my interpretation of the influence of the conventional origin of justice on the moral evaluation of justice by ‘us’, or people who live in civilized societies. In short, I will argue that the same form of general rules with which people recognise the ‘artificial’ interest of justice also directs the formation of their moral sentiments about justice. The most significant consequence of this influence is that when they morally approve of a just action or disapprove of an unjust action, they relate these actions *not only* to the character traits of the agents under concern *but also* to the ‘general scheme’ of actions regulated

Chapter 6
From Interest to Morality: Convention and Obligation

by the rules of justice. The relation to the rules of justice is, both concerning the meanings of just or unjust actions and concerning the ground of their moral values, more fundamental. Moreover, because it is the rules of justice which guide convention members' expectations about others' behaviour, rule-regulated external actions, instead of the character traits behind them, become the content of people's *idea* of the moral *obligation* to justice, even though their initial approbation or disapprobation is about the virtuous or vicious character traits of justice or injustice. Based on such special features of the moral sentiments about justice and the ideas derived from them, Hume's account of the origin of justice responds to the points that he makes about the 'artificial' features of our experience with the moral evaluation of justice.

Before starting, I need to acknowledge that though this line of explanation is developed on the basis of Hume's account of the origin of justice and his moral theory in general, it *goes beyond* what Hume provides explicitly in his theory of justice and its moral obligation. As we know, in the *Treatise*, Hume only offers *one* paragraph which specifically focuses on the moral obligation to justice (T 3.2.2.24; 499-500), and he gives no detailed explanation of how the artificial origin of justice can account for his observations in the Artificiality Arguments. To avoid exploiting the text, I will read this paragraph together with other relevant texts, and I take it as an interpretative principle that I shall avoid ascribing inconsistencies to texts as long as there is no compelling reason to do so. Also, I attempt to read Hume's philosophical texts in a way which makes them support themselves as much as possible. I will declare when my interpretation goes beyond what can be supported by textual evidence.

This chapter has three sections. I will start by explaining Hume's words that the ideas of 'justice', 'property', 'right', and 'obligation' all arise after the convention of property has been established (T 3.2.2.11; 490). The idea of obligation at stake here, as I will argue, refers to the *natural* obligation to justice and is based on the 'general sense of common interest' that people feel about observing the rules of justice. This sense is by itself 'artificial' because it arises from the custom-convention, and the idea of 'obligation', from the very time it was first formed, was specified in relation to the conventional rules of justice. In Section 6.2, I will argue against a widely accepted interpretation that the natural and moral obligations to justice are felt in two stages in the history of justice, namely, the former in nascent societies and the latter in civilized ones. On the contrary, I will argue that people in civilized societies preserve the perception of the sense of common interest though its motivational force wanes, and their sentiment of moral obligation to justice is generated *dependent on* the general rules with which they recognise this common interest. Finally, in Section 6.3, I will argue that because the moral sentiments about

justice in civilized societies are formed under the influence of convention-dependent general rules, just and unjust actions derive their meanings and moral values from the ‘general scheme of actions’ regulated by the rules of justice, and I will discuss how this dependence can account for the differences between natural and artificial virtues which I examined above in Chapter 4.

6.1 Natural Obligation, Artificial Interest: The Conventional Origin of Obligation

To relate Hume’s account of how the rules of justice are first established by conventions to his theory of the moral value and obligation of justice, the first key thesis which needs explanation is his claim that:

After this convention, concerning abstinence from the possessions of others, is enter’d into, and every one has acquir’d a stability in his possessions, there immediately arise the ideas of justice and injustice; as also those of *property*, *right*, and *obligation*. The latter are altogether unintelligible without first understanding the former. Our property is nothing but those goods, whose constant possession is establish’d by the laws of society; that is, by the laws of justice. Those, therefore, who make use of the words *property*, *right*, and *obligation*, before they have explain’d the origin of justice, or even make use of them in that explication, are guilty of a very gross fallacy, and can never reason upon any solid foundation. (T 3.2.2.11; 490-91)

Darwall (1995, 296–98, 314–15) argues that this passage is puzzling because neither of the two types of obligations Hume discusses in his theory, that is, natural and moral obligations, depends on convention to be intelligible. Darwall’s solution to this problem is that Hume means here a third type of ‘rule obligation’. While I have argued against this solution in previous chapters, the question Darwall raises here remains to be answered.

Another controversial issue related to this passage is how to understand Hume’s concept of ‘natural obligation’ when this phrase is opposed to ‘moral obligation’. To recall, Hume concludes the first part of his discussion of the origin of justice and begins the second one by remarking that the “*natural* obligation to justice, *viz.* interest, has been fully explain’d” and the moral obligation to justice remains to be accounted for (T 3.2.2.23; 498). The passage cited above appears immediately after Hume’s remark on the convention of justice and pages before he turns to the topic of moral obligation, which suggests that by this convention-dependent idea, he means the idea of *natural* obligation. Unfortunately, Hume does not explain what type of interest counts as a natural obligation. Moreover, he only talks about ‘natural obligation’ in his account of artificial virtues in the *Treatise* and provides no explicit remark on whether there is a natural obligation to have certain natural (i.e., non-artificial) virtues. In three places in the

6.1 Natural Obligation, Artificial Interest: The Conventional Origin of Obligation

Treatise and the second *Enquiry* where the topics do not specifically concern artificial virtues, Hume uses the term ‘obligation’ to address self-interest instead of morality. In Book 2 of the *Treatise*, he uses ‘such obligation’ to refer to “very strong motives of interest or safety” (T 2.1.10.6; 312) and talks about “strong obligations of interest to forbear any pleasure” (T 2.1.10.9; 314). In the *Enquiry*, he discusses whether there is an ‘interested obligation’ to be virtuous and explains this term with one’s “regard to his own happiness” (EPM 9.14; 278). Still, however, he does not explain in these texts whether an interested motive *as such* may be called an ‘obligation’ or it needs to meet some extra qualification to be sorted into this category, and he gives no clear indication of how these texts relate to the natural obligation to justice. Scholarly interpretations of Hume’s concept of natural obligation vary. Darwall reads Hume’s concept of natural obligation as inherited from Hutcheson who argues that one of the meanings of the term ‘obligation’ can be defined as “a Motive from Self-interest, sufficient to determine all those who duly consider it, and pursue their own Advantage wisely, to a certain Course of Actions” (*Inquiry* 2.7.1; 178). Gauthier (1992, 410) suggests that for Hume, wisely deliberated self-interest counts as a natural obligation especially when it contradicts the natural inclination of one’s interested desires. Don Garrett (2007, 267) argues that natural obligation “presumably lies in the tendency to feel non-moral prudential reproach for oneself for harming one’s own interests.” Neither the conflicts between interested desires nor negative self-evaluation in a prudential aspect, nevertheless, rely on conventions to be intelligible, and difficulties regarding the passage cited above still persist.

I defend the following reading in this section: (i) for Hume, the natural obligation to justice is constituted by the ‘general sense of common interest’ which he describes as people’s motive to establish the rules of justice in the account of agreement-convention; (ii) this sense of interest, considered from the perspective of the custom-convention, indeed arose gradually as a product of experience and custom; and (iii) the idea of obligation was first formed in relation to the conventional rules of justice and the ideas of ‘property’ and ‘right’, and it depends on the rules of justice to obtain its proper meaning. The interest to observe the rules of justice then makes up a first *exemplar* for impressions and ideas which fall under the general idea of ‘obligation’. It can be inferred from Hume’s account of interested and moral obligations that the extension of this idea later *expanded*, probably along different relations of similarity which gave rise to different subsets of this general category. The distinction between the first formation of the idea of obligation and its later application is helpful, especially in light of the discrepancies between Hume’s remarks on interested or moral obligations, examined by themselves, and his remarks

6.1 Natural Obligation, Artificial Interest: The Conventional Origin of Obligation

on ‘*property, right, and obligation*’, three ideas which he almost always examines in a group. Compared to moral obligation, there is less textual evidence to support a precise definition of ‘natural obligation’ as such, if Hume ever had such a definition in mind. I thus satisfy myself with an interpretation of Hume’s account of the natural obligation *to justice* without trying to settle the controversies over natural obligation *in general* in Hume’s theory.

The natural obligation to justice arises from the convention of justice because it is a sense of interest which people can only recognise in the ‘artificial manner’ that I discussed in the last chapter. This means that they think about the causal relation between just actions and their consequences through specific types of general rules with which they perceive such actions as instances of a ‘general scheme’ of similar actions performed by others in the entire community. Hume presents the form of such general rules in the model of agreement-convention in which every member of society finds observing the rules of justice conducive to their interest provided that others do the same. When they find such a sense of interest commonly shared in society, they form a resolution to pursue this interest. This account, as I argued above, cannot be taken as proof that in the ‘real’ history, the inventors of property are able to feel this sense of interest in this manner at the very temporal beginning of the development of justice. People in fact develop reflections on the interest of justice “insensibly and by degrees” (T 3.2.3.3; 503), as a consequence of the accumulation of experience of conflicts over external possessions and attempts to settle and avoid them. Their sense of interest attached to the observation of the rules of justice is then ‘artificial’ not only insofar as it depends on rational deliberation but also insofar as it is a *product* of the process which I call custom-convention, a product of various passions in social interactions together with the contribution of imagination and custom. In this reading, the following remark of Hume makes better sense:

[Justice and injustice have] two different foundations, *viz.*, that of *self-interest*, when men observe, that ’tis impossible to live in society without restraining themselves by certain rules ... ’Tis the voluntary convention and artifice of men, which *makes the first interest take place*. (T 3.2.6.11; 533, my italics)

The interest in observing property rules *arises from* “the voluntary convention of men”. If people needed such a sense of interest to invent such a convention, there would be a circle in the causal relation between the recognition of this interest and the establishment of this convention, but my interpretation avoids this problem. Also, Hume holds that “the interest, which gives rise to” the rules of property is not “of a kind that cou’d be pursu’d by the natural and inartificial passions of men”, and he explains this by the case that the rules of property

6.1 Natural Obligation, Artificial Interest: The Conventional Origin of Obligation

“seek their end in an oblique and indirect manner” because “their connexion with interest is somewhat singular”, that is, through a “whole plan or scheme” of actions (T 3.2.2.21-22; 497). The phrase “gives rise to” here, again, should be read not as implying temporal precedence but as suggesting that the rules of property are maintained by an ‘artificial interest’ the recognition of which, by itself, depends on the development of practices regulated by such rules in an earlier stage.

Let me dive into more details concerning the nature of this ‘artificial interest’ recognised in an ‘artificial manner’. In the agreement-convention, Hume writes the inventors of property express to each other a “general sense of common interest”. I take this *sense of interest*, which should refer to an impression, as constituting the content of the natural obligation to justice. A sense of interest *in abstaining from others’ possessions* should be an impression of reflection which arises from precedent impressions or ideas of such actions and makes one regard the latter as good to one’s interest. This sense of interest can be distinguished from one’s *desire* to pursue the recognised interest since while desire “arises from good and evil” (T 2.3.9.2; 438), not all perceptions about good or evil always give rise to desires. In this sense, the natural obligation to justice is not *by itself* a motive but an evaluative attitude toward just actions, though it can directly generate a motivating desire, and, because of this, there is no strict need to distinguish this sense of interest and its consequent motive. Finally, and most importantly, this perception concerns a ‘common interest’, which I understand as an interest that relates to each person’s self-interest in a *common* manner: “I observe, that it will be for *my interest* to leave another in the possession of his goods, provided he will act in the same manner with regard to me. He is sensible of a *like interest* in the regulation of his conduct” (T 3.2.2.10; 490, *my italics*). Everyone, according to this formula, finds a similar causal relation between just behaviour and their self-interest since the general scheme formed by all such mutually related actions serves as a means “by which we maintain society”, and society “is so necessary to their [our friends’] well-being and subsistence, as well as our own” (T 3.2.2.9; 489). This relation is “*singular*”: a just action, insofar as it is a *just* action, relates to one’s interest only as an instance of the general scheme of justice, and this general scheme relates to one’s interest only because it is necessary to maintain the peace and order of the society, and society is necessary for one to achieve various ends which are desirable due to one’s self-love or confined benevolence. The common interest, on this account, has a similar content for each person, namely, it concerns the maintenance of the order of the society. Though this means, ultimately speaking, obtains its value from each person’s ‘natural’ ends which are desirable by natural passions, its content

and relationship to these ends do not vary with the latter ends but have interpersonal and intrapersonal stability. To feel a natural obligation or sense of common interest in observing the rules of justice is to take such actions as good to oneself in this specific aspect, and this sense depends on the special general rules which arise from the custom-convention to be felt. In this aspect, my interpretation of Hume's remarks on the natural obligation to justice agrees most with that of Tito Magri (1996, 231–36), who also holds that this obligation consists in an 'artificial interest' the perception of which is shaped by convention-based general rules. Magri, meanwhile, suggests that Hume's concept of natural obligation may be extended to 'natural motives' other than self-interest when such motives are corrected by general rules. I am sceptical of this reading as I do not find it supported by sufficient textual evidence.

My reading has noteworthy implications in four aspects. First, I do not take this natural obligation equivalent to self-interest wisely calculated by instrumental reasoning, as the latter is not always guided by convention-based general rules. One can, as Hume addresses, in particular cases, find unjust actions conducive to one's interest and even the public interest, but this does not mean that in such cases, *the natural obligation* to justice fails, or one has a natural obligation to break the rules of justice. On the contrary, one may still feel a sense of interest in *observing* the rules of justice, which makes up a feeling of natural obligation. Second, this natural obligation is usually a negative evaluative feeling. As one's perception of this sense of interest arises from one's experience of inconveniences arising from people's transgression of property rules (T 3.2.2.10; 490), and as maintaining the order of the society makes up a necessary condition for one's well-being but does not contribute much to one's well-being directly, one should more usually get aware of the interest of justice by *painful* feelings when *unjust* actions produce harmful consequences and disappoint people's expectations about others' behaviour. Third, because of this second feature, and because the code of behaviour required by the natural obligation to justice is characterised in a negative manner, namely, in terms of *abstaining* from others' property, actions which *conform with* the rules of justice and, in this sense, discharge one's obligation to justice do not always need to be *motivated* by the sense of common interest. When one is cultivating or reaping fruits from one's own field and, by these very actions, *not* trespassing on one's neighbour's property, for instance, one is acting in conformity with property rules, but one's motive for doing so may simply be self-love and love for one's family instead of a sense of natural obligation. The latter sense comes to the fore usually only when one is tempted to violate property rules or have done so. In this sense, the natural obligation to justice works to *redirect* one's natural interested desires but does not make

up the *redirected* interested desires. Redirected self-interest takes ‘gratifying all our appetites’ as its end (T 3.2.2.13; 492), while the natural obligation only functions to restrain one’s natural desires instead of positing any positive ends for one to pursue. Finally, as a consequence of the three former points, in this stage, when one develops a stable *character trait* to observe property rules, this trait cannot be characterised by the ability to act on wise instrumental reasoning or the natural obligation to justice taken as a motive at work whenever one acts in conformity to property rules. Rather, such a person should (i) be familiar with property rules, namely, be able to recognise what objects and actions pertain to others’ ‘property’ and ‘property right’; (ii) be able to feel the sense of common interest in a proper manner, i.e., with the help of convention-based customary general rules; and (iii) act in conformity with property rules and, *if* tempted to breach such rules, can feel the natural obligation to justice in such strength as to be able to restrain such desires. This analysis will provide a first model for the virtue of justice in Section 6.3. Now, however, let me first say more about natural obligation in this section.

As the sense of the natural obligation to justice, being an impression of reflection, relies on conventions to be felt, it is evident that the *idea* of this natural obligation, which is copied from this impression, depends on conventions to be formed. My reading also makes it easy to explain why the idea of this obligation is usually discussed together with the ideas of ‘property’ and ‘right’, and why all three ideas rely on the ideas of justice and injustice to be intelligible (T 3.2.2.11; 490-91). To make sense of Hume’s remarks in this paragraph, with which I began my discussions in this section, it is important to distinguish between two orders, the temporal order of the *generation* of people’s perceptions about the convention of property, and the order of *explanation* in which one settles the meaning of one idea by appealing to other ideas. It is clear that by ‘justice’ in this passage, Hume means “the *laws* of justice” (T 3.2.2.11; 491, my italics) or the set of rules about which part of others’ external possessions one should abstain from. As moral judgements remain to be examined at this point in Hume’s investigation, the ‘should’ here indicates a natural obligation: people would respond to actions which breach such rules with a negative feeling since it harms the common interest. Concerning the order of *generation* in the history of justice, the sense of common interest, which constitutes the natural obligation to justice and the basis for the idea of this obligation, arises gradually together with the abstract ideas concerning property, property rights, the rules of justice, and qualities of actions and agents which are characterised in terms of justice (say, a just action is an action which conforms with the rules of justice, and a just person has a disposition that I specified above). All these perceptions arise gradually from people’s shared experience about conflicts

over external possessions and attempts to avoid them, and they make up people's customary cognitive and evaluative pattern, a pattern which gradually becomes habitual to them and can be enlivened and used in both pre-reflective evaluations and actions and reflections about justice. In this process, the sense of common interest works as a crucial motive which supports the practice of property, and only when people gradually become aware of this interest and impose mutual constraint on each other does the code of behaviour in abstaining from others' property obtain the status as *rules*. In this sense, it can be said that the rules of justice depend on this natural obligation to be maintained.

Meanwhile, concerning the order of *explanation*, the impression and idea of obligation precede the ideas of property and right, and the ideas concerning the rules of justice precede that of obligation. To recall, in the first argument in *Treatise* 3.2.6, Hume defines property as an "internal relation" between an external relation between a person and an object, on the one hand, and this external relation's 'influence on the mind' on the other (T 3.2.6.3; 527). While the internal influence he discusses in this argument is a sense of duty, when the convention of property is first invented, the sense of common interest can serve the same function. One's property, in this stage, can be explained as those parts of one's external possessions which make all other people feel a natural obligation to abstain from, and one can also derive from this relation the idea of property right.¹²² All these ideas rely on the rules of justice to be intelligible because the natural obligation to justice is by itself an 'artificial' interest: one perceives this sense of interest when one conceives single actions to abstain from a certain part of others' possessions as instances of the general scheme of just actions, or as actions which agree with the rules of justice. The natural obligation to justice, on this account, is an obligation to observe the *rules* of justice; one can explain the obligations and rights about property with the rules of justice, but one cannot explain 'justice' with 'right' or 'obligation'. Consequently, "[t]he origin of justice *explains* that of property. The same artifice *gives rise to both*" (T 3.2.2.11; 491, my italics). Both the ideas of justice and property arise from convention, and they likely emerge and evolve in the same process. To explain the origin of the idea of property (and right and obligation), meanwhile, one needs to resort to the origin of the rules of justice.

¹²² Based on Hume's account of the work of imagination in the production of the 'moral relation' of property in *Treatise* 3.2.3.4n (504), 3.2.3.7 (506), and 3.2.3.10n (510), his definition of property in *Treatise* 2.1.10.1 (309-301), and his note on the definition of property and property right in *A Dissertation on the Passions* (DP 2.30n; 13), I suggest that the 'moral relation' of property is a fiction produced by imagination in analogy to the causal relation of possession. The idea of property right, correspondingly, is a fictitious quality of a proprietor formed in analogy to the idea of causal power. A more detailed discussion of this topic goes beyond the scope of this thesis.

6.2 Custom and Sentiment: From Natural to Moral Obligation

Let me sum up. I argued in this section that the ‘natural’ or non-moral obligation to justice is ‘non-natural’ or artificial: it is constituted by a sense of common interest in avoiding conflicts over external possessions and maintaining the order of society, and people’s feeling of it relies on convention-dependent general rules which relate just actions to each person’s interest only through a ‘general scheme’ of similar action. This explains Hume’s suggestion that the idea of obligation only becomes intelligible after the convention of property has been established if we consider the natural obligation to justice as a first exemplar of the general idea of ‘obligation’. Based on this view, I also argued that the natural obligation to justice differs from both wisely calculated self-interest, which may characterise the virtue of prudence, and redirected interest, which more usually makes up people’s motive for actions which conform with property rules. The impression and idea of obligation, together with the ideas of property, and right, make up people’s customary cognitive and evaluative pattern which gradually becomes habitual to them as the convention of justice develops. In the next section, I will investigate how this habit may influence the moral evaluation of justice.

6.2 Custom and Sentiment: From Natural to Moral Obligation

In my examination of Hume’s account of the conventional origin of the rules of justice and the natural obligation to justice from Section 5.2 to Section 6.1, I refrained from discussing moral sentiments or sentiment-based moral judgements. In this section, I bring the latter topics back to the story by discussing the relationship between the natural and moral obligations to justice. As I observed in Section 5.1, many scholars maintain that these two types of obligations are at work in different stages in the history of justice: the former in small and nascent societies, and the latter in large and civilized ones. I argue in this section that this interpretation is imprecise. In my reading of Hume, people in large societies can still feel the sense of natural obligation to justice though its motivational force wanes. Then, I discuss how the same cognitive and evaluative habit which makes people recognise the interest of justice influences the formation of the sentiment of the moral obligation to justice.

The central text which supports the reading that the natural and moral obligations can be separated into two historical stages goes as follows:

To the imposition then, and observance of these rules [concerning the stability of property], both in general, and in every particular instance, they are at first mov’d only by a regard to interest; and this motive, on the first formation of society, is sufficiently strong and forcible. But when society has become numerous, and has encreas’d to a tribe or nation, this interest is more remote; nor do

6.2 Custom and Sentiment: From Natural to Moral Obligation

men so readily perceive, that disorder and confusion follow upon every breach of these rules, as in a more narrow and contracted society. But tho' in our own actions we may frequently lose sight of that interest, which we have in maintaining order, and may follow a lesser and more present interest, we never fail to observe the prejudice we receive, either mediately or immediately, from the injustice of others; as not being in that case either blinded by passion, or byass'd by any contrary temptation. (T 3.2.2.24; 499)

This reading usually accompanies the view that natural and moral obligations are, by definition, two types of motives. In large societies, natural obligation quits from people's motivational set since the interest of justice becomes more difficult to perceive and less 'strong and forcible'. It then leaves a vacancy in the motives which can support the convention of property and people's disposition to observe property rules, and the sentiment of moral obligation, or desires derived from this sentiment, fills up this vacancy. I take this line of interpretation to be imprecise since first, as I argued in Chapters 1, 3 and this chapter, I do not read Hume as defining either the natural or the moral obligation as a determining motive. Both, instead, are reflective impressions which express some (usually negative) evaluative attitudes. The natural obligation to justice is made up of a sense of common interest attached to ideas of just or unjust actions, which can, in general, produce motivating desires to pursue this interest or avoid harming it. It is then one question whether this *sense* of interest can be perceived by people in large societies, and another question whether the *motive* it produces can be strong enough to guarantee stable observation of property rules in such a condition.

Hume does not deny that people in civilized societies can feel a sense of common interest related to just or unjust actions. He argues that the negative consequences of unjust actions on social order become "more remote" when societies become larger, but he also remarks that this interest is "nevertheless *real* for being remote" (T 3.2.7.3; 535, my italics). Also, as he suggests in the second *Enquiry*, though people are often accustomed to observing the rules of justice mechanically and without reflecting on the first motive to do so, it is still true that "even in common life we have every moment recourse to the principle of public utility, and ask, *What must become of the world, if such practices prevail? How could society subsist under such disorder?*" (EPM 3.47; 203). The sense of interest in justice, for people in large and civilized societies, then, is 'remote' but not inaccessible. Hume explains this point by arguing:

[A]s all men are, in some degree, subject to the same weakness, it necessarily happens, that the violation of equity must become very frequent in society, and the commerce of men, by that means, be render'd very dangerous and uncertain. You have the same propension, that I have ... You are,

6.2 Custom and Sentiment: From Natural to Moral Obligation

therefore, naturally carry'd to commit acts of injustice as well as I. Your example both pushes me forward in this way by imitation, and also affords me a new reason for any breach of equity, by showing me, that I shou'd be the cully of my integrity, if I alone shou'd impose on myself a severe restraint amidst the licentiousness of others. (T 3.2.7.3; 535)

It appears that to recognise the harm of injustice, one needs to *imagine* the universal violation of the rules of justice and infer how this would endanger “the commerce of men”. This, however, appears contrary to his claim that the interest of justice is “real”. For Hume’s claim about this ‘reality’ to be warranted, in my view, this line of inference must be, in a certain sense, based on one’s *real* experience, but to what extent can this condition be granted?

Two points may be made here. First, in my view, one important reason why the relation between just or unjust actions and the common interest is still ‘real’ for people in large societies is that the circumstances which necessitated the invention of justice do not change in civilized societies: the natural inclinations of human beings remain selfish, partial, and jealous, and the provision of external goods remains limited compared to the needs arising from people’s avidity. As long as these inclinations are unchanged, people’s passions and interests will keep clashing with each other in social interactions, and both the temptation of benefiting from breaching the rules of justice and the need for mutual constraints to guard against such breaches will persist. As long as these interactions, which, as I argued in the last chapter, make up the fundamental layer of Hume’s account of the development of justice, remain unaltered, the reflections which emerge from such experiences will keep being made in the structure that Hume presents in the agreement-convention, and such reflections will keep generating habitual mental associations in ways similar to what supposedly happened to the inventors of justice. It remains true for ‘us’, therefore, that single unjust actions tend to be imitated if not prevented by mutual restraints among convention members, and some interest commonly shared by all convention members would then be endangered. No individual in large societies, of course, needs to experience the invention of the practices of property and promises or that of the abstract ideas of ‘property, right, and obligation’. ‘We’ are born into such conventions and learn the rules of justice and the meaning of these terms by education, together with, as Hume remarks, teachings about why observing such rules is useful to others and ourselves. At the same time, ‘we’ *also* experience various forms of injustice and their harmful consequences in various scales, or at least can have such experiences in principle, with which our ideas of the interest of justice may be enlivened not only by the ‘artificial cause’ of education but also by the ‘natural cause’ of “custom and repetition” (T 1.3.9.19; 117). In this sense, Hume’s account of the origin of justice also gives

an analysis of the structure of the social interactions and individual agents' psychological development by which the practice of justice is maintained, although he does not clarify the relationship between the genealogical and analytical functions of his account.¹²³

Second, if the uniformity of human nature in different historical stages and forms of society can grant that the experiences, reflections, and cognitive and evaluative habits about the interest of justice can keep being regenerated in large societies, it must be noted that they may be similar to the supposed experiences of the inventors of justice in structure but different in many other aspects. One may suggest that the most significant difference here is that when justice was first invented, “disorder and confusion” actually “follow upon every breach of these rules”, while this does not apply to larger societies. As I argued in Section 5.3, however, people’s recognition of the relationship between just or unjust actions to the common interest, when they invent the rules of justice, should already be generalized through the ‘general scheme’ of similar actions. The most significant difference here, in my view, rather consists in the *degree* of generalization in which one conceives just or unjust actions as instances of a ‘general scheme’. In smaller societies where the scope of social cooperation does not extend beyond people’s acquaintances, people may have vivid experiences of the mutual constraint among convention members, and they can easily resort to such experiences to derive ideas about the ‘general scheme’ of justice and the interest it secures. As societies expand, the ‘general scheme’ made up of just or unjust actions of all convention members also exceeds the scope of personal experiences and becomes an object of imagination concerning anonymous agents and evaluators. Accordingly, it is more difficult to “*readily perceive*” how single breaches of the rules of justice relate to “disorder and confusion” in society (T 3.2.2.24; 499, my italics). Although people in larger societies may experience various forms of injustice and their harmful consequences on various scales, if they are to generalize such experiences on a societal scale, then, they have to make more efforts in imagination and inference. The sense of common interest, consequently, will be less vivid and strong for them than for people in smaller societies. This is how Hume explains his use of the word “remote”: he uses it in a figurative sense, not to address a temporal or spatial distance or the length of the causal chain from an unjust action to its harmful consequence, but to address

¹²³ As Harris (2020, 85) observes, Hume does not make it clear in the *Treatise* whether his account of the origin of justice “describes something that happened very early in human history or a learning process that each and every one of us has to go through in order to be made fit for adult social life”. These processes must be different, however, as Harris (2019, 234) argues in another paper, because ‘we’ do not need to go through the conjectural process in which property and promises are first invented to have a concept of justice. This problem may be partly solved if ‘our’ experiences about the consequences of unjust actions have some structural similarity with Hume’s account of the conjectural origin of justice, but it is true that Hume does not provide a clear and fully developed account of this point.

the strength and liveliness of perceptions, since objects closer in time and space also have “a proportional effect on the will and passions” (T 3.2.7.2; 534-35). This ‘remoteness’, or the lack of vivacity and strength of the sense of common interest due to a higher degree of generalization, is what distinguishes people’s conception of the interest of justice in larger societies from that in smaller societies. Because such a generalizing process is based on, or can be based on, real experiences of the harm of injustice, this conception is “nevertheless *real* for being remote”.

As long as the common interest secured by the rules of justice can be ‘really’ recognised by people in large and civilized societies, though it becomes too ‘remote’ to provide a constant motive to justice, it is still true that people can perceive a sense of natural obligation to justice. The reason why Hume emphasizes the change from nascent societies and civilized ones at the beginning of his account of the moral obligation to justice, then, is not that the natural and moral obligations are felt in two different historical stages, but that the moral obligation to justice plays a more indispensable role in maintaining this convention in civilized societies. Hume does not address exactly at which time in the history of justice people start to feel moral sentiments about justice. He says that moral sentiments about justice arise *after* the invention of the rules of justice (T 3.2.5.12, 3.2.6.11; 523, 533). As the establishment of the conventions of justice indeed takes place in a gradual process, and as people can already sympathize with victims of unjust actions in this process, there is no strong reason to hold that moral sentiments about justice arise after the conventions of justice have been established with a considerable degree of stability or expanded to a considerably large scope. It is more plausible that moral sentiments arise gradually in the same process as the conventions of justice develop, though conditioned by people’s recognition of the interest of justice. I hereby turn to the second topic of this section, namely, how people’s sense of natural obligation to justice and the mental habit underlying this sense influence the moral evaluation of justice and injustice.

Hume’s words on the formation of moral sentiments about justice are quite brief:

[W]e never fail to observe the prejudice we receive, either mediately or immediately, from the injustice of others; as not being in that case either blinded by passion, or byass’d by any contrary temptation. Nay when the injustice is so distant from us, as no way to affect our interest, it still displeases us; because we consider it as prejudicial to human society, and pernicious to every one that approaches the person guilty of it. We partake of their uneasiness by *sympathy* ... And tho’ this sense, in the present case, be deriv’d only from contemplating the actions of others, yet we fail not to extend it even to our own actions. The *general rule* reaches beyond those instances, from which

it arose; while at the same time we naturally *sympathize* with others in the sentiments they entertain of us. (T 3.2.2.24; 499)

Though Hume does not deny that we morally *approve* of justice and regard it as a *virtue*, his principal case in this account is how we morally *disapprove* of injustice, which sentiment, as I argued in Chapter 3, makes up the sentiment of moral obligation. I will likewise primarily focus on disapprobation in my discussion.

The sentiment of the moral obligation to justice appears to arise in the following manner as it is delineated in the passage above: we, as spectators, observe an unjust action and its harmful consequences and sympathize with the uneasiness of its victims. This sympathetic pain makes up a ‘sense of moral evil’ and makes us call this action and the character trait associated with it ‘vicious’. Then, we follow general rules to extend this sentiment to our own actions and disapprove of our actual or potential failure to act justly. This account, however, stays in need of amendment based on Hume’s discussion of the generation of moral sentiments in *Treatise* 3.3.1, as he himself reminds his readers of (T 3.2.2.23; 498). One crucial topic that Hume does not address in this passage is that qualified moral judgements are made from “some *steady* and *general* points of view” (T 3.3.1.15) so that spectators in different relationships with the agents who are evaluated can agree on their judgements. Adopting a general point of view at least partly consists in neglecting one’s particular relationship to the evaluated agent and restricting the objects of one’s sympathy to those “who have any commerce with the person we consider” (T 3.3.1.18; 583). To do so, one follows *general rules* about the influence of traits of character and makes evaluations based on “general views of things” (T 3.3.1.23; 587).¹²⁴ It follows from Hume’s later account that *general rules* are involved in the formation of the moral obligation to justice not only when we extend our disapprobation of others’ unjust actions to ourselves but also in the previous step when we first feel the sentiment of disapprobation. To disapprove of

¹²⁴ Hume’s account concerning the ‘general points of view’ in the *Treatise* is unclear in some significant respects, and many controversial issues arise in scholarly interpretations of this part of his moral theory. This thesis does not aim at settling all these controversies, though some necessary clarifications are needed here. First, I agree with Cohon (2008b, 132) and Don Garrett (forthcoming) that there are two different issues with which Hume deals in this part of his discourse: the first is that when making moral judgements, we neglect the evaluated agents’ influence on our own *interest*, and the second is that we *also* neglect our *connections* to the evaluated agents and those who commerce with them which result in variations in the strength of our sympathetic feelings. Although Cohon and Garrett hold that adopting the general points of view only concerns the latter issue, I take it as concerning both. This is because Hume clearly addresses both issues in relevant texts (for instance, T 3.3.1.17-18, 3.3.1.30; 582-84, 591), though he does not draw a clear-cut distinction between them. Second, I agree with Baier (2006, 114) and Ben-Moshe (2020, 443, 451) that the point of view that one adopts to evaluate a specific trait of character may vary according to the type of trait under concern. Third, I largely agree with Chamberlain (2022, 82–90) that general rules are at work in the formation of all moral sentiments so that evaluative sentiments and passions which are not formed through general rules are not ‘*moral*’.

such actions *morally*, we need to consider them from a ‘general point of view’ and sympathize with those who are or would be influenced by such actions in general. Whom, then, do we sympathize with when disapproving unjust actions, and what type of loss do we consider as pertinent when partaking in their uneasiness?

Many scholars hold that the object of sympathy relevant to the moral evaluation of justice and injustice is *the whole society*, and to conduct such evaluation, we consider the increase or loss of public interest, understood as the interest of the whole society which comprehends each individual’s self-interest, in the long run. They support this interpretation with Hume’s remark that “a sympathy with public interest is the source of the moral approbation, which attends that virtue” (T 3.2.2.24; 499-500).¹²⁵ There are reasons to doubt that this view, if not untenable, at least needs significant modification. First, as Haakonssen (1981, 35–36) insightfully observes, Hume’s theory of sympathy requires that “we can only have sympathy with specifiable individuals” as only then can the *similarity* between the ideas of their persons and our own persons induce the transformation of ideas of others’ feelings into their correspondent impressions. The ‘public’ whose interest is served by justice, nevertheless, is “a group with a changing membership of non-specific persons”, and its anonymity makes it unsuitable as an object of sympathy; also, it is too demanding for ordinary evaluators to take concern with and calculate the public interest in the long run. Second, Hume does not distinguish the original motive for the invention of justice and the ground of people’s moral approval of justice in some texts. For instance, when summarizing his entire account of the three ‘fundamental laws of nature’, he remarks that:

[W]e are to consider this distinction betwixt justice and injustice, as having two different foundations, viz. that of *self-interest*, when men observe, that ’tis impossible to live in society without restraining themselves by certain rules; and that of *morality*, when this interest is once observ’d to be common to all mankind, and men receive a pleasure from the view of such actions as tend to the peace of society, and an uneasiness from such as are contrary to it. (T 3.2.6.11; 533)

Morality, or the second foundation of justice, arises when people observe ‘*this interest*’ to be commonly shared by all mankind, and ‘this interest’ refers to the *self-interest* which attends the maintenance of the order of the society. It has the same content, in other words, as the natural

¹²⁵ See, for instance, the readings of Gauthier (1979, 16–18), Haakonssen (1981, 35–36), Sugden (2005, 175–76), and Cohon (2008, 149). Hume makes similar remarks in other texts like *Treatise* 3.3.1.9, 3.3.1.12; 577, 579.

obligation to justice, and an apparent difficulty arises in reconciling this remark with Hume's words in *Treatise* 3.2.2.24.

Both issues, in my view, will become less problematic if the moral disapprobation of unjust actions is formed under the influence of the same general rules which relate just or unjust actions to the 'artificial' common interest. This means that when observing an action of injustice, one considers it as an instance of a 'general scheme' of similar actions which lead to the loss of a common interest, and one considers this interest shared by all convention members since everyone expects this interest to be secured by all fellow people's observation of the rules of justice. The victims of this unjust action, then, include "*every one* that approaches the person guilty of it" (T 3.2.2.24; 499, my italics); it is true that this action is harmful to society, but it is harmful to society *insofar as* it fails every society member's expectation about others' just conducts and the preservation of the common interest. To sympathize with the common interest, in this sense, is to sympathize with a group of individual persons, though an indefinite group represented by some specifiable exemplars. Moreover, the relationship between the harm on these exemplars and the harm on the indefinite group of victims is not between part and whole but between particular and general, and the mental process involved here is generalization rather than adding up individual society members' gain and loss. Also importantly, the same generalizing process can make the evaluators recognise their *similarity* to the objects of their sympathy since all are members of the same social convention and subjects of the same types of rights and obligations, which contributes to the operation of sympathy. Concerning the sympathy which promotes "the natural obligation to honesty into a moral obligation", Baier (2010, 47) argues that "the only sort of sympathy ... there can be is sympathy with right-holders, with victims of theft or fraud, and those who fear it". I agree with Baier, and this may respond to Haakonssen's worry to a certain extent.

This reading also resolves the apparent discrepancy between Hume's words in *Treatise* 3.2.2.24 and 3.2.6.11. If the ground of the sentiment of moral disapprobation is the unjust action's harm to the *common* interest, it makes sense to say both that one sympathizes with the *public interest* and that this ground is the same as the *self-interest* which people observe in abiding by the rules of justice. This is because the common interest both makes up a part of each society member's self-interest and, insofar as it is commonly shared by all people, makes up a part of the public interest, though its extension is narrower than 'public interest' considered

in the aggregational sense.¹²⁶ Public and private are rather two perspectives to conceive the common interest. In the first part of his account of the origin of justice, Hume frequently describes the interest of justice by juxtaposing these two perspectives: justice is “requisite to the public interest, and to that of each individual”; its rules are established out of “a concern for our own, and the public interest”; the general scheme of justice is “absolutely requisite, both to the support of society, and the well-being of every individual”; and it “is infinitely advantageous to the whole, and to every part” (T 3.2.2.20-22; 496-98). In my view, it is in the same sense that he says when disapproving an unjust action, we consider it as “prejudicial to human society, and pernicious to every one that approaches the person guilty of it” (T 3.2.2.24; 499).

It follows from this reading that the relationship between the natural and moral obligations to justice is much closer than suggested by many commentators. The same general rules which make people feel the natural obligation also direct them to the objects to sympathize with, and the sense of common interest, which makes up the natural obligation to justice, also accounts for the sympathetic pain which makes up the sentiment of moral obligation. Later, when Hume examines the topics of promises and allegiance, he argues that their natural obligations “give rise to” or “cause” their moral obligations (T 3.2.8.7, 3.2.9.3; 545-46, 551).¹²⁷ Also, regarding promissory obligations, he claims that after people have recognised the interested obligation to perform their promises, “a sentiment of morals *concur*s with interest, and becomes a new obligation upon mankind” (T 3.2.5.12; 523, my italics), which suggests that these two feelings may exist and exert their forces together. I regard the same relationship as applicable to the natural and moral obligations to justice in the narrowest sense. In short, the natural obligation to justice causes its moral obligation, and this is true both to the development of the practice of justice on a social and historical level and to individual, psychological processes.

Before moving on to the next topic, I would like to respond to two worries which may be raised about my interpretation. The first is that interpreting ‘public interest’ in Hume’s account of the moral obligation to justice as the common interest cannot respond to Haakonssen’s worry regarding the anonymity of the object of sympathy since “every one that approaches the person

¹²⁶ At this point, I agree with Sugden’s (2009, 16) reading that when Hume argues that the moral approbation of justice arise from a sympathy with the public interest, he “is building some notion of common interest into what he calls ‘public interest’”.

¹²⁷ Importantly, the natural obligation to allegiance is only a *partial* cause of its moral obligation since other factors also contribute to this psychological process, which makes it happen that when the interest to obey a government ceases to exist, the sentiment of moral obligation may still be in force (T 3.2.9.3; 551). I think the same is true for the obligations to justice.

guilty of it” is still too general and indefinite an idea for the mechanism of sympathy, especially in large societies. I think it is true that this difficulty cannot be entirely resolved by my reading. One plausible further response to this concern may be developed by focusing on a significant omission in the *Treatise* account of the disapproval of injustice, that is, Hume curiously keeps silent regarding the *individuals* who are directly harmed by unjust actions in this passage, while in experience, these people’s feelings may more usually be a source of strong sympathetic pain. Hume acknowledges the importance of this part of the ground of the moral disapprobation of injustice in the second *Enquiry*, where he observes that:

[A]fter the laws of justice are fixed by views of general utility, the injury, the hardship, the harm, which result to any individual from a violation from them, enter very much into consideration, and are a great source of that universal blame which attends every wrong or iniquity ... by depriving me of it [my property], you disappoint my expectations, and doubly displease me, and offend every bystander. It is a public wrong, so far as the rules of equity are violated: it is a private harm, so far as an individual is injured ... What injures the community, without hurting any individual, is often more lightly thought of. But where the greatest public wrong is also conjoined with a considerable private one, no wonder the highest disapprobation attends so iniquitous a behaviour. (EPM App.3.11; 310-11)

What Hume distinguishes here, and what he does not distinguish in the *Treatise*, is that unjust actions usually have two different groups of victims: those directly injured by such actions and “every bystander” whose expectation of others’ just actions is disappointed. It is plausible that the sympathetic pain one feels about the “private harm” is usually more vivid than that felt about the “public wrong”, and when these two feelings are conjoined, the former may work to enliven the latter and make up for the difficulty to perceive it vividly due to the abstractness of the idea of an average bystander. If we substitute the model concerning the object of sympathy in *Treatise* 3.2.2.24 with the twofold model in the second *Enquiry*, at the same time, the specific form of general rules which relate unjust actions to the common interest remain indispensable for the generation of moral disapprobation of injustice since spectators, as discussed in this passage in the *Enquiry*, rely on such general rules to recognise injustice as a “public wrong” and to relate this aspect to its harm to private individuals and make these painful feelings

conjoin. This passage, then, can supply the *Treatise* account but does not necessitate a substantial change in my interpretation of the latter.¹²⁸

The second worry is that if in large societies, the natural and moral obligations to justice “concur” with each other, this would make moral obligation in such societies less important. This worry may arise especially because the sentiment of moral obligation cannot function as a stable motive for justice at an initial stage without the ‘second artifice’ of education. It has no significant advantage, in other words, compared to the sense of natural obligation regarding its initial motivational force. Setting aside their motivational force and capacity to be moulded by education (though the latter may play a significant role here), the sentiment of moral obligation has an advantage that the natural obligation to justice, constituted by the common interest *taken as* a part of one’s own interest, does not have, that is, the sympathetic pain that one shares with victims of other people’s unjust actions is less prone to be “either blinded by passion, or by ass’d by any contrary temptation” (T 3.2.2.24; 499). Apart from the increase of the scope of social cooperation and the corresponding degree of generalization required when one thinks about the general scheme of justice and the common interest promoted by this scheme, another crucial difference between small and large societies emphasized by Hume is that “the pleasures of life” increase in large societies, and the lively perceptions of immediate enjoyment may more easily attract one’s attention and blind one’s sense of common interest (T 3.2.8.1; 539). I argued above that the common interest can be conceived from both private and public perspectives. When it is conceived as a part of one’s self-interest, it competes with *other* interested desires to occupy one’s attention and determine one’s will and action. When it is conceived from a spectatorial perspective, as an interest shared by all others, meanwhile, it does not need to be involved in such competitions. For this reason, and also because one sympathizes not only with the victims’ pain but also with other spectators’ similar moral disapprobation, the common interest becomes more ‘visible’ to average convention members. Its relevant pain and pleasure are more stably felt from a spectatorial instead of an agential perspective, and its relevant ideas become more frequently enlivened. Even when moral sentiments have not gained sufficient motivational force from the second artifice, then, it is crucial that they enable one to maintain lively perceptions about the harm and benefit of injustice and justice in a way that self-interested

¹²⁸ One reason Hume might have for not highlighting the sympathy with the particular victims of unjust actions in the *Treatise* is that, as he addresses in the Artificiality Arguments, he intends to explain why we blame unjust actions even when they, taken by themselves, do not bring harm to the parties directly involved (T 3.2.1.13; 3.2.6.9, 482, 532). In the second *Enquiry*, likewise, he does not mention this ground of moral disapproval until the very end of his account of justice.

considerations cannot. The first significant influence that the expansion of societies has on the practice of justice, then, in my view, is that it requires such a switch of perspective: in large societies, one has to first reach out to observe, imagine, and sympathize with the pain caused by injustice and then *turn back* to evaluate one's own actions to overcome the bias caused by the increasing pleasures of life. This may explain why Hume emphasizes on the distinction between self-interest and the public interest when commenting on the ground of moral evaluation of justice in *Treatise* 3.2.2.24, though, in my reading, 'public interest' here stands for the common interest conceived from other people's perspectives.

6.3 Law, Obligation, and Virtue: Replies to the Artificiality Arguments

I have by far discussed one crucial aspect in which the moral sentiments 'we' feel about justice in large and civilized societies are 'artificial', namely, the formation of such moral sentiments is guided by convention-based general rules with which we regard every single action of justice or injustice as an instance of a 'general scheme' of similar actions and relate it to the common interest in preserving the order of the society. Insofar as this common interest also constitutes the content of the natural obligation to justice, it makes sense to claim, just like Hume claims regarding allegiance, that "the natural obligation of interest" is the cause of "the moral obligation of duty" (T 3.2.9.3; 551). Importantly, this reliance on convention does not contradict Hume's emphases on the point that the moral approval of justice arises "*naturally*" (T 3.2.2.25, 3.2.6.11, 3.3.1.12; 500, 533, 579). This is because, in the latter texts, the naturalness of moral sentiments is contrasted to the second artifice of education: such sentiments are 'natural' because they arise from the natural psychological mechanism of sympathy, instead of being crafted entirely out of non-moral passions like pride. At the same time, such sentiments are 'artificial' in light of their reliance on the first artifice of convention, which Hume also emphasizes in many places of the *Treatise* (T 3.1.1.10, 3.2.1.1, 3.2.1.17; 475, 477, 483). This part of Hume's overarching explanatory project now reaches its end.

By finishing his account of the origin of justice at this point, however, Hume leaves many questions unanswered, especially concerning how the artificiality of the 'sense of justice' can account for the discrepancies between this sentiment and the natural inclination of moral sentiments that he addresses in the Artificiality Arguments in *Treatise* 3.2.1 and 3.2.6. As I discussed in Chapter 4, Hume claims in these arguments that (i) we regard the laws of justice inflexible and insensitive to particular circumstances; (ii) we regard the rights and obligations of justice insusceptible to an insensible gradation; (iii) we define property with an internal sense

of duty felt about certain actions and objects instead of physical connections between persons and objects, and (iv) we regard the sense of duty as a sufficient motive to observe the rules of justice without requiring to explain this moral motive by a non-moral motive. Though Hume suggests that these ‘artificial’ features of our moral evaluation of justice can be explained by its conventional origin, it is not immediately clear how his answer can match his questions. I argued in Chapter 4 that one common reason for these ‘artificial’ features is that regarding justice, it appears that we evaluate actions as right or wrong according to their relationship to a set of moral *laws*. By the natural inclination of our moral sentiments, on the other hand, we primarily react to people’s *character* and judge it as virtuous or vicious. For Hume’s theory of justice to reply to the Artificiality Arguments, then, it must be able to explain how moral sentiments, which are naturally attached to ideas of *persons*, can operate in a way that is regulated by *impersonal* laws which forbid and permit regardless of the characters and situations of the persons under concern (T 3.2.6.9; 532). Unfortunately, to a considerable extent, Hume leaves this task to his readers. To investigate *to what extent* Hume’s moral theory can respond to this question, I turn to his long story about convention and obligation on which I have been working in the four previous sections. My following discussions in this chapter, compared to previous ones, will be supported by less textual evidence and rely more heavily on my past conclusions. I start by discussing the relationship between the virtue of justice and the laws of justice according to Hume’s theory of the conventional origin of both. Because the laws of justice, as I will argue, maintain a certain independence in people’s conceptions from individual convention members’ character traits, when they think about the moral obligation to justice, external rule-regulated actions take precedence over character traits to make up the content of the ideas of such moral obligations. This feature, though not explicitly addressed by Hume, may account for the ‘artificial’ characteristics of the evaluation of justice that he observes in his arguments.

One crucial aspect in which Hume’s genealogical account about justice can help to explain the moral phenomena he observes in the Artificiality Arguments is that this account sheds light on the relationship between the moral rules or laws of justice and the types of character traits which are described in terms of ‘justice’ and ‘injustice’. I observed in Chapter 4 that in *Treatise* 3.2.6, Hume claims both that the notions of property, right, and obligation depend on the *virtue* of justice (T 3.2.6.3; 527) and that these notions are determined by the *rules* of justice (T 3.2.6.6; 528). His account of the conventional origin of justice helps to clarify the relationship between these two claims because this account has two implications about the ontological status of the

rules or laws (Hume does not distinguish these two terms when discussing the origin of justice) of justice. First, at least before (or abstracted from) the invention of governments and civil laws, the laws of justice only exist as embodied in the convention members' disposition to recognise a certain part of others' external possession as their property, to abstain from them, and to feel a sense of natural and/or moral obligation to do so when one is tempted to do otherwise. They do not, in other words, exist as imposed by a lawgiver and enforced by the latter's power; they exist in convention members' codes of behaviour and are maintained by the mutual constraint of these people who are motivated by their both natural and artificial senses of interest. I argued in Section 6.1 that in earlier stages of the development of justice, a person can be described as having the mental quality of *justice* if they have obtained a convention-dependent cognitive and evaluative habit which is partly expressed by the habit to feel a sense of natural obligation and to act on it under proper circumstances. Because moral sentiments, as I argued above, may be involved in the history of justice from its earlier phases on, and because this mental quality is useful both to others and to the agent under concern, this type of trait qualifies to be regarded as *virtuous*. Now, because in the order of *generation*, as I discussed in Section 6.1, the rules of justice depend on convention members' sense of natural obligation and their disposition to feel it, it makes sense to say that the rules of justice depend on the virtuous character of justice.

This, meanwhile, is only one side of the story. The second implication of the conventional origin of justice is that the structure of this convention, that is, single actions of justice can only produce their expected benefit by the uniform concurrence of the majority of the whole society, requires the rules of justice to be *formulated* and *observed* with strong intra- and inter-personal stability and uniformity. This means that people's conception of such rules and expectations of others' observation of them, especially after the convention of justice become stable and widely accepted enough, are and should be (as required by the common interest) specified in a manner that is *independent* of the particular characters and circumstances of the agents at issue. When Hume explains the inflexibility of the laws of justice with their conventional origin, he argues that when people "have agreed to restrain themselves by general rules", they at the same time agreed that such rules should be "unchangeable by spite and favour, and by particular views of private or public interest" (T 3.2.6.9; 532). This is because they had in their view not only the inconvenience caused by people's natural avidity but also that caused by their natural partiality in evaluating people's actions and characters. The latter threatens the peace of human society no less than the former, and the common interest requires that the rules of justice cannot be applied "differently in every particular case, according to every particular utility" (T 3.2.3.2;

502). The “general nature of the question”, in the application of the rules of justice, should take an overwhelming weight, while in other cases of moral evaluation, general tendencies are taken into consideration as well as particular situations (T 3.2.6.9; 532). As a consequence, people’s conception of the rules of justice maintains a certain independence from particular convention members’ character traits. Everyone is expected to adopt a certain code of behaviour, no matter whether they are wise or foolish, benevolent or selfish, having a strong mind or not, having a strong concern for doing their part in a mutually beneficial scheme or not, because what is most necessary for the common interest to be secured is the uniformity in their rule-regulated actions. Though in the order of generation, the rules of justice depend on an earlier form of the virtue of justice to come into being and be maintained, in the order of *explanation* or concerning the way people conceive relevant ideas, especially after the convention of justice has been widely and stably established, the rules of justice gradually gain an independent status as universal and inflexible laws and, as I argued in Section 6.1, define the contents of the ideas of property, right, and obligation. This responds to the first point in Hume’s observations on the artificiality of justice as I listed above, namely, concerning the inflexibility of the laws of justice.

Now consider how this relationship between the virtue and rules of justice influences the moral evaluation of *actions* which fall under the general ideas of ‘justice’ and ‘injustice’. The idea of an action which, for instance, abstains from a certain part of others’ external possessions can be further specified both in relation to the agent’s motives and mental dispositions and in relation to conventional rules about property. When we morally approve of such an action, or, as it may more usually happen, when we morally disapprove of one’s failure to do so, according to the natural inclination of moral sentiments, we regard this action or omission as expressing or signifying a certain aspect of the agent’s character. This is, in my reading of Hume’s theory, *psychologically necessary* for the generation of moral sentiments, and it is considering this association that we call this agent and their action ‘virtuous’ or ‘vicious’. At the same time, because we feel this sentiment of moral approbation or disapprobation, as I argued in the last section, by following convention-dependent general rules, we *also* regard this action as one which observes or breaks the rules of justice. This latter association is also necessary for the generation of moral sentiments because just actions, insofar as they are *just*, are only related to the ground of sympathetic pleasure or pain *as* instances of the general scheme of actions (T 3.3.1.12; 579). Importantly, the former ‘natural’ relation between actions and characters relies on the latter ‘artificial’ relation between actions and rules to produce the relevant sentiments of approbation or disapprobation, because a single individual’s just or unjust character trait, like

a single action of justice or injustice, has to depend on a society-wide coordinational scheme to be considered as useful or harmful. In other words, to judge an action as virtuous or vicious with respect to justice or injustice, it is inadequate to look beyond the external action to consider its underlying motive and character; one also needs to look beyond the motive and character to consider the general scheme of actions regulated by the rules of justice.

As this general scheme provides a necessary background for the moral *evaluation* of the virtue of justice and the vice of injustice, it likewise plays an indispensable role in the *characterisation* of the latter qualities: such qualities cannot be properly specified without referring to the rules of justice because if they were specified in another manner, they would not be useful or harmful to the society and to oneself in the same aspect. The virtue of justice, in its earlier form, is one's disposition to have regard to others' *property* and *right* and to feel and act on a sense of natural *obligation* in certain situations, and the vice of injustice should plausibly be featured by one's lack of such regard.¹²⁹ As the ideas involved in this specification depend on the rules of justice to be properly defined, so is the idea of the virtue of justice and the vice of injustice. The moral evaluation of just or unjust actions as 'virtuous' or 'vicious', then, from the very beginning, is shaped by the rule-regulated practice of justice which keeps certain independence from individual convention members' character, and to recall, it is with an undeclared premise that the virtue of justice can only be characterised in relation to the rules of justice that Hume argues in the second part of the Artificiality Argument in *Treatise* 3.2.1 that neither extensive nor confined benevolence can be the original motive to justice.

As a final move, I consider the formation of the sentiment of moral obligation to justice and the idea derived from it, namely, what one is morally obliged to in this respect. I argued above in Chapter 3 that for Hume, the sentiment of moral obligation is a special type of disapprobation which arises from the idea concerning the absence of an action or quality of the mind (T 3.2.5.4; 517). It can arise from two possible routes: when people expect a virtue but this expectation is failed, and when people disapprove of a vice and associate it with an idea of its remedy. As one feels such a disapprobation, one can derive from it the idea that one lies under the moral obligation to have the absent mental quality or to perform the neglected action. In the case of justice, the sentiment of moral obligation can arise in both ways: Hume's account in *Treatise* 3.2.2.24 takes the route of the disapprobation of vice when we, as spectators, pay

¹²⁹ Hume argues that there is 'neither justice nor injustice' in the state of nature previous to the invention of property (T 3.2.2.28; 501). Though people in such a state could be evaluated as vicious by their unchecked avidity, partiality, and cruelty, then, they could not have the vice of *injustice*.

attention to the harm done by unjust actions, while the structure of the convention of justice indicates that every convention member expects all others to have the disposition to act as the conventional rules require. Though the causes of the sentiment of moral obligation, in my view, are ideas of the *vice* of injustice or the lack of the *virtue* of justice, both being ideas of character traits, the convention-based general rules underlying the formation of this sentiment direct the evaluator's attention to *actions* which observe the rules of justice and make such actions the content of the *idea* of moral obligation or duty of justice. This is because it is *primarily* such rule-regulated actions that people *expect* from other convention members and regard as remedies for the harm caused by injustice; as the idea of moral obligation is about the content of such expectations, there is no need to reach further than rule-regulated actions to form such an idea.

Let me dive into more details. By 'primarily' here, I mean both that rule-regulated actions are conceptually more *fundamental* when people form their ideas about this moral obligation and that such actions are practically more *important* so that they attract more attention.¹³⁰ From the conceptual perspective, when one derives from the disapprobation of an unjust action the idea of moral obligation to justice, because one relies on the rules of justice to explain what the vice of injustice or the expected virtue of justice is, one does not need to resort to this vice or virtue to think about what people generally *need* or *expect for* in this respect. To see this, imagine two people in a conversation: if one asks the other 'What should I do?' or 'What should this person do in this situation?', the other may directly answer 'Do as the rules of justice require' instead of 'do as a just person would do', because to do as a just person would do *is* simply to do as the rules of justice require.¹³¹ This does not mean that the character trait of justice cannot mean more than a mere disposition to perform actions in conformity with the rules of justice. This trait, as a mental quality, rather denotes a specific cognitive, evaluative, and conative habit, a habit to pay proper regard to others' property rights and to feel and act on a sense of natural obligation (still, I only consider the earlier form of the virtue of justice in this stage) when tempted to violate others' property right. Insofar as this is a quality which all society members expect others to have, it is also true that everyone in the society has a moral obligation to have the virtuous quality of justice. One does not, however, need to resort to this obligation for the virtue of justice to explain their obligation to perform just actions. On the

¹³⁰ I use 'conceptual' to address the structure of the relationship between impressions and ideas, thus in line with Hume's psychological approach to the formation and meaning of ideas.

¹³¹ I benefit from Margaret Watkins Tate (2005, 105–11) for this imaginary conversation.

contrary, because the virtue of justice obtains its moral value on the ground of the benefit of the rule-regulated general scheme of justice, one rather relies on the obligation to observe the rules of justice to explain the obligation to have the virtue of justice.

Second, from the practical perspective, since the stability of the convention of justice and the security of the common interest require the rules of justice be observed by the majority of society members with a strong interpersonal uniformity, each person's specific motivational status occupies a lesser place in their mutual expectation about each other's conduct compared to the external rule-conformity of their actions. While one earlier form of the virtue of justice can be characterised by the sense of natural obligation, each individual's psychological makeup in this respect may diverge with, for instance, their ability to reflect on the common interest and their strength of mind. As one's sense of common interest may be stronger or weaker, one's observation of the rules of justice may be more or less constant. The mutual expectation among convention members about just actions, meanwhile, remains strict and uniform because it is on this uniformity that the common interest depends. Also, as I argued in Section 6.1, because the requirements of the rules of justice are negative, i.e., to abstain from doing something, the extension of actions *in conformity with* the rules of justice is much wider than that of actions motivated by the sense of natural obligation. This also makes external actions, instead of their motives, occupy people's focus when they form ideas about the moral obligation to justice.

Importantly, the same priority of external actions compared to character traits *cannot* apply to judgements about moral obligation when it comes to natural virtues. Hume also talks about the distribution of attention when one thinks about one's moral obligation in the passage on the duty of gratitude. An ungrateful person, when observing his lack of a virtue which is common in human nature, disapproves of his character and hates himself on this account. He may then regard mere grateful actions as a remedy for this deficiency and be motivated by a regard of moral obligation to perform such external actions. In this case, Hume argues, "[a]ctions are at first only consider'd as signs of motives: But 'tis usual ... to fix our attention on the signs, and neglect, in some measure, the thing signify'd" (T 3.2.1.8; 479). How does this differ from the formation of the idea of the moral obligation to justice that I discussed above? They differ because when just or unjust actions are morally evaluated, they are not merely taken as *signs* of just or unjust traits of character; they obtain their meanings and moral values also from the rule-regulated general scheme of actions. Naturally virtuous actions do not rely on such a general scheme to be characterised and morally approved. They pursue some 'natural good' which is desirable to human beings' natural passions, and they are recognised as useful or

agreeable based on such natural desirability, according to, as I argued in Chapter 5, Hume's more considered approach to the distinction between natural and artificial virtues (T 3.3.1.12; 579). Accordingly, one does not need to refer to 'the concurrence of mankind' when forming general rules about the useful or agreeable tendency of natural virtues, and there is comparably less need to expect all people to observe some strict, precise, and uniform codes of behaviour. External actions of gratitude or other natural virtues, then, obtain their moral merit directly from the virtuous qualities of the mind which typically cause them, and when one expects others or oneself to have such virtuous qualities, the expectation about such qualities takes precedence over the expectation about their correspondent actions both concerning the explanation of the content of such expectation and concerning their practical significance. What one ought to do with regard to gratitude, then, is not only to act as a grateful person acts but also to feel as a grateful person feels; the external action only signifies the trait of character. If one fixes one's attention on the sign and neglects the thing signified and, upon performing grateful actions, thinks that one has fully discharged one's moral obligation in this respect, one makes a *mistake* as one's character remains deficient. In the case of justice, meanwhile, since just actions are not mere signs of a just character, and since rule-regulated actions take precedence in people's expectations, if one observes the rules of justice out of a sense of duty, one does fully discharge one's moral obligation in this respect. There is no need, in other words, to keep feeling self-mortified and struggling to "acquire by practice, that virtuous principle, or at least, to disguise to himself ... his want of it" (T 3.2.1.8; 479).

Now, as long as people can derive from their sentiment of moral obligation to justice the idea of this obligation which has just actions as its content, this sentiment and its corresponding idea can 'concur' with the sense of natural obligation and its corresponding idea and become a part of people's habitual cognitive and evaluative patterns about property and promises. Moral obligation can then join with natural obligation to define property so that in later stages of the history of justice, we can say that property consists of "some internal relation" which external person-object relations cause on the mind "by giving us a sense of duty in abstaining from that object, and in resorting it to the first possessor" (T 3.2.6.3; 527), and the idea of moral obligation can then become a part of the group of associated ideas of 'property, right, and obligation'. Interestingly, it is possible that, from a genealogical perspective, the idea that people derive from the disapprobation of injustice and associate with their expectations for just actions falls under the general category of '*obligation*', because this idea assimilates the idea of the natural obligation to justice which, as I suggested in Section 6.1, served as an initial

exemplar for the abstract idea of ‘obligation’. These two ideas are similar since both represent painful evaluative attitudes towards the violation of the rules of justice, and both are related to expectations for strictly rule-abiding actions. These similarities exist because, as I argued in Section 6.2, the same convention-dependent associative habit guides the formation of both the sense of natural obligation and the sentiment of moral obligation to justice, and the former sense also partly causes the latter sentiment through people’s sympathy with some pain which exemplifies the loss of the common interest. It is possible, though purely out of conjecture, that it is in this causal relation and structural similarity where the uniformity between the natural and moral obligations to justice lies: the idea of obligation first arises in the form of natural obligation and is then applied to the idea derived from the moral disapprobation of injustice.

Finally, to see how this entire account may respond to the second and fourth points related to the virtue of justice, we need to see how this virtue can be characterised in later stages in the history of justice, namely, in large and civilized societies when the natural obligation to justice becomes too weak to provide a constant motive for just actions. Though, as Hume argued, the sentiment of moral obligation may by itself be “too weak to control our passions” (T 3.2.2.24; 500), after being magnified by public and private education and being closely connected to the regard to one’s character and reputation, this sentiment, or motives generated by this sentiment, can become strong enough to confront temptations and motivate at least some society members to act justly. It can then become a motive for justice and constitute a part of the psychological makeup of a just person, and the virtue of justice, at this stage, can be characterised by one’s disposition to pay proper regard to others’ property rights and to feel and act on a sense of duty in certain situations. Again, depending on one’s endowments and education, this moral motive may be stronger or weaker, and one’s disposition to act on it can be more or less reliable. The virtue of justice, in this sense, admits of degrees. A person with a strong sense of duty may “fix an inviolable law to himself, never, by any temptation, to be induc’d to violate those principle” (T 3.2.2.27; 501). An average member of a well-organized society with government and civil laws, meanwhile, may observe the rules of justice and disapprove of injustice habitually, but they may not be able to resist strong temptations with their sense of duty and rather rely on the fear of punishment to abide by property rules. The former is virtuous in this respect; since the latter’s character trait is also useful to society, it is also virtuous, but only to a very moderate

6.4 Conclusion

extent.¹³² Concerning this feature, then, the *virtue* of justice (and the *vice* of injustice) does not diverge from “all kinds of vice and virtue” which “run insensibly into each other” (T 3.2.6.7; 529). This does not, however, make the *rights* and moral *obligations* of justice admit of insensible degrees because the latter ideas derive their meanings and moral value from the laws of justice. Finally, because in their civilized condition, almost every society member makes moral evaluations about justice in the manner shaped by convention and education, they should all have formed or been inculcated with the sentiment of moral obligation to justice and its idea in the manner as I discussed above. They will not, correspondingly, like the supposed person in his “rude and more *natural* condition” (T 3.2.1.9; 479), regard just actions as deriving their moral value from a non-moral but virtuous or obligatory motive. On the contrary, they will take the answer of the “honest man”, that is, “my regard to justice, and abhorrence of villainy and knavery” is a sufficient motive to act justly, as “just and satisfactory” (T 3.2.1.9; 479) without requiring for more explanation. By explaining the social, historical, and psychological causes underlying this attitude of satisfaction, Hume’s theory of justice then responds to the concern raised in his first Artificiality Argument.

6.4 Conclusion

I discussed in this chapter how the specific cognitive and evaluative habit, which emerged from the process in which the practice of justice gradually developed through conventions and keeps being regenerated and maintained by the dynamics of passions, imagination, custom and reason in social interaction, gives rise to the impressions and ideas of the natural and moral obligations to justice. I argued that the natural obligation to justice is an ‘artificial’ sense of interest: it is constituted by the general sense of common interest in maintaining the order of the society, and this sense of interest is ‘artificial’ because it can only be related to single just or unjust actions by convention-dependent general rules by which one recognise such actions as instances of a ‘general scheme’ of similar actions. From this sense of interest, people derive their first idea of ‘obligation’ which correlates with the ideas of ‘property’ and ‘right’ and depends on the rules of justice to be intelligible. Then, I suggested that the same convention-dependent general rules also direct the formation of moral sentiments about justice, and the ‘public interest’ that people

¹³² On this point, I agree with Watkins Tate (2005, 110–16), Baier (2010, 32, 68, 110), and Harris (2010a, 40–41) that members of a civilized society usually have various motives to act justly, and this is consistent with Hume’s account of the moral merit of justice. Harris draws on this point to argue that Hume rejects the idea that a particular motive or disposition may identify a “truly just person”, but I think this point is consistent with Cohon’s (2008b, 171–82) view that justice as a virtuous character trait is characterised by the sense of duty, as long as the virtue of justice admits of degrees.

6.4 Conclusion

sympathize with when disapproving of an unjust action is the same as the general sense of common interest, though it is conceived from the perspectives of other members of society instead of in relation to one's self-interest. As Hume argues, the natural obligation is a partial cause of the moral obligation. Finally, because of this influence of convention-dependent general rules in the formation of moral sentiments, just or unjust actions are morally evaluated not only in reference to characters but also in reference to the laws of justice. Because the laws of justice maintain a certain degree of independence from individual agents' characters in people's conception, and because just or unjust actions more fundamentally rely on the general scheme of rule-regulated actions to obtain their meanings and moral values, such actions take priority when people form ideas about the moral obligation to justice.

When developing this line of explanation, I did not discuss *to what extent* the convention-dependent habits and people's reflections on the general scheme of justice can counteract their natural inclinations to pursue present and particular interests and to focus on them when making moral evaluations. Hume's position on this point, in my view, is moderate: convention shapes our moral sentiments and sentiment-based judgements about justice to some extent, but it exerts its influence together with a group of other factors, especially the 'second artifice' of education. I did not pay much attention to education in my discussion not because it is unimportant but because I would like to investigate how much the 'first artifice' of convention could do its job in Hume's explanatory project without having to rely on the 'second artifice', as Hume himself makes it clear that he hopes to distance his theory of justice from a Mandevillian account in which the moral value and motive of just are merely products of the artifice of politicians and parents (T 3.2.2.25; 500). It is probably impossible, and also unnecessary for Hume's explanatory project to decide on the accurate degrees to which these two artifices contribute to the formation of 'our' moral conceptions and opinions about justice, and I leave this as an open question here.

Chapter 7

Conclusion: Morality between Nature and Artifice

In her paper ‘Modern Moral Philosophy’, Elizabeth Anscombe suggests that “the concepts of obligation, and duty ... ought to be jettisoned if this is psychologically possible; because they are survivals, or derivatives from survivals, from an earlier conception of ethics which no longer generally survives, and are only harmful without it” (1958, 1). This situation of the concept of obligation, according to Anscombe, is discovered by Hume: as Hume’s time was one in which “the belief in divine law had long since be abandoned”, his distinction between ‘is’ and ‘ought’ reveals an interesting situation “of the survival of a concept outside of the framework of thought that made it a really intelligible one” (1958, 6). Although doubts may be shed on Anscombe’s characterisation of Hume’s time and her interpretation of Hume’s ‘is/ought’ distinction, I think her statement about the situation of the concept of obligation, especially *moral* obligation, does describe the situation which confronts Hume, who takes over this concept from previous moral theories and needs to make it intelligible in a new conceptual framework in which divine laws have no place. Hume would not think, in my reading of him, that jettisoning this concept would be ‘psychologically possible’, or maintaining it in his moral theory would be ‘harmful’. He gives it a distinctive place in his virtue-centred moral theory, and this thesis is an attempt to work out how he does so, especially in the *Treatise*, by his distinction between ‘natural’ and ‘artificial’ virtues and his account of the social and historical account of the origin of the rules of justice and the ‘natural’ and ‘moral’ obligations to observe such rules. In this concluding chapter, I will first briefly summarize the line of argument I made in this thesis and then make two pieces of extensive remarks about the implication of my interpretation: the first concerns Hume’s view about the moral obligation to have certain natural virtues, the second concerns the functions of Hume’s account of obligation apart from descriptive and explanatory ones.

The first two chapters of this thesis were dedicated to locating Hume’s account of obligation in its context, regarding both its place in the framework of his moral theory and its relationship to previous theories of obligation. Hume does not have a systematic theory of obligation, and probably he did not intend to develop such a theory. An important feature of

Hume's conception of morality is that a considerably large part of it, including many morally good or bad actions and characters and the evaluation of them, can, at least in principle, do away with the concept of obligation. We approve or disapprove of them as what is virtuous or vicious, praiseworthy and blameworthy; these feelings and ideas constitute a crucial part of our self-conception and relationship to others through the indirect passions of pride and humility, love and hatred, which, in turn, makes up a crucial reason why virtues are appealing and encouraged in the social world, and vices, unappealing and discouraged. Hume does not make obligation a central concept in this basic description of our moral practice partly because he shares Hutcheson's positive view on human nature's self-sufficiency in leading a morally good life and partly because this is how, based on his anatomy of human nature, moral sentiments work. His interest in the topic of obligation, considered against this general background, is driven more by questions confronted by this picture of morality than by an interest in this concept itself. Two problems which Hume takes under concern are that the natural inclinations of human passions and evaluative faculties are not completely sufficient to make them sociable, on which he agrees with Mandeville, and that moral judgements about rights and obligations, especially about perfect and adventitious rights and obligations, are significantly different from those about virtues and vices, on which he agrees with Grotius and Pufendorf. Hume also benefits from these authors in his approach to solving these problems, namely, by his account of two 'artifices', convention and education, through which the conceptions of justice, right, and obligation arise and work to secure human sociability. By comparing Hume's distinctions between 'natural' and 'non-natural' obligations and taxonomies of obligation in natural law theories, I argued that he substitutes the traditional framework of divine legislation with a framework of the development of social order and moral conceptions to account for the relationship between human nature and their moral obligations.

Chapters 3 and 4 turn to Hume's basic delineation of the sentiment of moral obligation, its features as an evaluative attitude, its origin, objects, and probable scope, as well as its different characteristics when applied to justice and to 'natural virtues' like gratitude and benevolence. I observed that for Hume, the sentiment of moral obligation is a specific type of disapprobation, it arises when the lack of a virtue is regarded as vicious or when the disapprobation of a vice is associated with the idea of its remedy, its objects, or the things which are regarded as obligatory based on this sentiment, are, in many cases, mental qualities signified by motives, abilities, or habits, no matter whether such qualities can be acquired voluntarily or not. However, because Hume does not pay much attention to the moral obligation to have certain natural virtues, and

because his account of obligation is so closely intertwined with his account of artificial virtues that neither of these two topics could be well-examined without diving deeply into the other, however, this general observation concerning the feature and origin of moral obligation cannot but be sketchy and conjectural in many respects. What Hume makes much clearer is that the obligation to have certain natural virtues is *different* from the obligation to justice. Judgements on the former are typically imprecise when applied to particular actions and sensitive to particular circumstances and characters of the people involved. This is the case because, as I read Hume, they are derived from antecedent conceptions of virtues and general expectations about character traits directed by considerations of the common course of human nature, one's social roles, and probably also character exemplars, which are usually neither strict nor universally applicable to all people in all circumstances. The obligation to justice, on the other side, together with the notions of right and property, is judged in reference to strict and universal laws or rules of action, and as such laws are in tension with the natural inclinations of human passions and evaluative faculties, the virtue of justice, if characterised as one's disposition to observe such laws, cannot, at the same time, be characterised by any natural motive. I argued that it is the law-directed characteristics of the obligation to justice which motivate Hume's distinction between natural and artificial virtues in the *Treatise* and raises the requirement for a genealogical account of justice as a part of Hume's overarching project in this work to explain the causal principles underlying our sentiment-based moral evaluations.

Finally, Chapters 5 and 6 investigate Hume's genealogical account of justice and discuss to what extent it can work to explain the law-directed characteristics of the obligation to justice. What Hume provides in the *Treatise* is a complicated account with several aspects, layers, and functions, and he does not distinguish them clearly. It addresses the development of some social institutions as well as individual person's moral conceptions acquired within such institutions, and the latter is, furthermore, analysed in terms of Hume's psychology of mental associations. It describes some conjectural events in the remote past of human societies, but it is also meant to, in my reading, analyse the form of social coordination regulated by the rules of justice and explain how this form of social coordination shapes the way in which people evaluate the utility of justice and the moral values of just or unjust actions and characters, and his discussions of the latter two aspects are meant to be applicable to all times and societies as long as human beings' natural inclinations and circumstances with regard to the acquisition and distribution of external goods remain uniform. One key observation which helps to clarify the relationship between these different aspects and concerns is the distinction between the custom-convention

and agreement-convention, according to which Hume provides *one* fundamental account of the gradual development of the institution of justice, motivated by the clash of selfish or partially benevolent passions and made possible by the contribution of imagination, reason, and custom, and *another* derivative account of people's rational reflections on the common interest served by this institution. The latter account emerges from the former and captures some structural features of the institution of justice and the customary general rules with which people conceive and evaluate just actions; both accounts work not only to explain the past of justice but also, to a certain extent, analyse its present. The gist of this entire account is that the apparently 'non-natural' characteristics of the laws and obligation of justice exist because human nature makes them so: the distinction between 'mine' and 'thine' is regulated by universal rules, evaluated in reference to a mutually expected interest of all convention members, and generates ideas of obligations about external actions rather than motives or character traits, because all these are required for human beings to live a sociable life. Although this point can be made both concerning the historical origin of justice and the situation of the institution of justice in large and civilized societies, of course, it is also important that the 'sense of common interest' in observing the rules of justice is much weaker and dependent on a process of imaginative generalization or universalization in civilized societies, which makes governments, civil laws, education, and concern about honour and reputation, all being later phases of human artifice, essential to maintaining the practice of justice in such societies.

With all these discussions made, there are still two questions about Hume's conception of moral obligation which remain to be answered. The first is whether Hume's account of the origin of the 'artificial' obligation to justice can shed any light on the obligation to have certain natural virtues like humanity and gratitude. As I noted in this thesis, Hume not only claims that the obligation to justice depends on conventions but also claims that the very *idea* of 'obligation' depends on the convention of justice to be intelligible (T 3.2.2.11; 490-91). I suggested in Chapter 6 that the first *idea* of obligation which people have, in Hume's genealogical account, is the idea of the natural obligation or the common interest in abstaining from others' property. It is possible, although merely out of my conjecture in line with Hume's psychology of cognition, that as moral sentiments later 'concur' with this sense of interest, and as the idea derived from the disapprobation of injustice share similarities with the idea of the natural obligation to justice, the scope of the general idea of obligation was then extended to the former, moral idea, and hence arises the idea of the 'moral obligation' to justice. If this conjecture may be granted, would the idea that 'one lies under an obligation' to have a certain virtue, or behave

as people with a certain virtue would behave, arise because of some similarity between the idea derived from the disapprobation of one's lack of this virtue and the idea of the moral obligation to justice? The idea of the moral obligation to have a natural virtue like gratitude, then, arises when we apply the artificial idea of 'obligation' to a natural virtue, probably in an 'imperfect' manner as natural virtues are unlike justice in being regulated by precise and universal laws. Shaver (1992, 550) makes a similar suggestion: it is probable that "[t]he idea of obligation as a limit on our natural propensities arises with property. Once it arises, we can look back and describe ourselves as being obliged to have the natural virtues before property". This piece of conjecture has crucial implications for Hume's conception of moral obligation in general. The idea of the obligation to justice would then influence judgements about the obligation to have certain natural virtues as an archetype, and the latter judgements would then be influenced by considerations about the necessity of the virtue under concern to the subsistence of the society, about a mutual interest which one may expect in return for performing one's duty, and so on.¹³³ The same influence would also make all judgements about obligation take actions rather than character traits as their objects and make the scope of obligation restricted to what can be voluntarily performed or acquired. This line of thought, however, can only remain a piece of conjecture.

The second question is whether Hume's account of obligation, apart from its descriptive and explanatory functions which I have made the main subject issue of my discussions above, also has other functions. I noted in Chapter 1 that Hume does not mean his theory of morality to serve a directly practical purpose, and it is controversial whether various themes in his theory, like the category of virtue and the 'general point of view' from which moral judgements are made, are merely descriptive or also normative as well. It is also controversial what normative implications his moral theory may have: I observed that scholars like Hursthouse (1999) and Swanton (2015) read Hume as a virtue ethicist; at the same time, especially concerning Hume's theory of justice, scholars like Harrison (1981) and Chwaszcza (2013) read it as a version of rule-consequentialism, while Gauthier (1979) and Sugden (2009) read it as contractarian. Two brief points may be made here. First, generally speaking, I think Hume's explanatory approach to moral studies has an important critical function: it generates criticisms against some moral

¹³³ Both possibilities are also mentioned by Shaver (1992, 547–53). In 'Of Suicide', Hume claims that "[a]ll our obligations to do good to society seem to imply something reciprocal" (E-Su; 586). This cannot be explained by his accounts of the obligation to have social virtues like humanity and gratitude in the *Treatise* since the naturally virtuous motives of humanity and gratitude do not take with them a requirement for reciprocity, but this can be explained if 'all our obligation to do good to society' has some similarity with the obligation to justice.

opinions, principles, and social plans when they cannot be supported by his analyses of the social and psychological mechanisms underlying human beings' moral practices, which can be exemplified by his criticism against some schemes of the distribution of external goods in both the *Treatise* (T 3.2.3.2; 502) and the second *Enquiry* (EPM 3.23-25, 192-94), his critical points concerning many opinions on virtues and vices like those on pride and 'monkish virtues', and many of his moral and political essays. Hume's account of obligation, in my reading, also has such a function, and one way in which this function may be substantiated is that his account of moral obligation sheds some doubt on all the normative theories mentioned above while, at the same time, also incorporates some aspects of them as a part of 'our' opinions about obligation. This is the second point I would like to make here: I think Hume's account of moral obligation is not virtue ethical because regarding justice, as I discussed above, considerations about law, obligation, and right precede those about motives and character traits, but regarding the natural virtues, the evaluation of virtue and vice precedes that of obligation. Hume's theory of justice may appear rule-consequentialist as single actions of justice are evaluated in reference to the interest served by the rule, or the entire rule-regulated general scheme of just actions. However, it diverges from rule-consequentialism because, as I noted above, the moral obligation to justice arises from sympathy with the common interest rather than the collective interest of the whole society, and this common interest, which lies in securing human sociability, is not an object for maximization. The role the common interest plays in the conventional origin of justice and its obligations may support contractarian readings of Hume to a certain extent, but the recognition of this interest cannot directly serve as a ground of justification in Hume's sentimentalist theory, and I also observed above that Hume disagrees with social contract theories which appeal to a supposedly real explicit or tacit contract, nor would he agree with hypothetical contractarianism. Moreover, no matter how similar Hume's account of 'artificial' moral obligation might appear to rule-consequentialism or contractarianism, his account of natural virtue does not align with these theories. All these considerations provide another reason why the distinction between so-called 'natural' and 'artificial' domains and aspects of morality is important for Hume. To return to the point with which I started this thesis, I agree with Baier (2010, 124) that Hume accounts for moral obligation in his virtue-centred theory in a "nonreductive distinction-preserving way", and I also agree with Wiggins (1998, 140) that Hume's distinction between natural and artificial virtues addresses subdomains of morality which are both "mutually irreducible" and "mutually indispensable"; this, in my view, is where the attractiveness of Hume's theory of obligation lies.

Bibliography

Primary Sources

- Allestree, Richard. 1700. *The Whole Duty of Man Laid down in a Plain and Familiar Way for the Use of All*. London: W. Norton for Edward Pawlet.
- Cumberland, Richard. 2005. *A Treatise of the Laws of Nature*. Edited by Jon Parkin. Translated by John Maxwell. Indianapolis: Liberty Fund.
- Grotius, Hugo. 2005. *The Rights of War and Peace*. Edited by Richard Tuck. Translated by John Morrice. 3 vols. Indianapolis: Liberty Fund.
- Hobbes, Thomas. 1994. *Leviathan*. Edited by Edwin Curley. Indianapolis and Cambridge: Hackett Publishing Company.
- . 1997. *On the Citizen*. Edited by Richard Tuck and Michael Silverthorne. New York: Cambridge University Press.
- Hume, David. 1975. *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. Edited by L. A. Selby-Bigge. Revised by P. H. Nidditch. 3rd ed. Oxford: Clarendon Press.
- . 1978. *A Treatise of Human Nature*. Edited by L. A. Selby-Bigge. Revised by P. H. Nidditch. 2nd ed. Oxford: Clarendon Press.
- . 1987. *Essays Moral, Political, and Literary*. Edited by Eugene F. Miller. Indianapolis: Liberty Fund.
- . 1998. *An Enquiry Concerning the Principles of Morals: A Critical Edition*. Edited by Tom L. Beauchamp. Oxford: Clarendon Press.
- . 2007a. *A Dissertation on the Passions and The Natural History of Religion*. Edited by Tom L. Beauchamp. Oxford: Clarendon Press.
- . 2007b. *A Treatise of Human Nature: A Critical Edition*. Edited by David Fate Norton and Mary J. Norton. 2 vols. Oxford: Clarendon Press.
- . 2011. *The Letters of David Hume*. Edited by J. Y. T. Greig. 2 vols. New York: Oxford University Press.

Bibliography

- Hutcheson, Francis. 2002. *An Essay on the Nature and Conduct of the Passions and Affections, with Illustrations on the Moral Sense*. Edited by Aaron Garrett. Indianapolis: Liberty Fund.
- . 2007. *Philosophiae Moralis Institutio Compendiaria, with A Short Introduction to Moral Philosophy*. Edited by Luigi Turco. Indianapolis: Liberty Fund.
- . 2008. *An Inquiry into the Original of Our Ideas of Beauty and Virtue*. Edited by Wolfgang Leidhold. 2nd ed. Indianapolis: Liberty Fund.
- . 2014. *A System of Moral Philosophy*. 3 vols. Cambridge Library Collection Philosophy. Cambridge: Cambridge University Press.
- Justinian, Caesar Flavius. 1913. *The Institutes of Justinian*. Translated by J. B. Moyle. 5th ed. Oxford: Clarendon Press.
- Kames, Henry Home Lord. 2005. *Essays on the Principles of Morality and Natural Religion*. Edited by Marie Catherine Moran. Indianapolis: Liberty Fund.
- Mandeville, Bernard. 1988. *The Fable of the Bees, or Private Vice, Publick Benefits*. Edited by F. B. Kaye. 2 vols. Indianapolis: Liberty Fund.
- Pufendorf, Samuel. 1729. *Of the Law of Nature and Nations*. Translated by Basil Kennett. 4th ed. London: J. Walthoe, R. Wilkin, J. and J. Bonwicke, S. Birt, T. Ward, and T. Osborne.
- . 1744. *De Jure Naturae et Gentium, Libri Octo: Cum Integris Commentariis Virorum Clarissimorum*. Francofurti & Lipsiae: Ex Officina Knochiana.
- . 1991. *On the Duty of Man and Citizen According to Natural Law*. Edited by James Tully. Translated by Michael Silverthorne. Cambridge: Cambridge University Press.
- . 1994. *The Political Writings of Samuel Pufendorf*. Edited by Craig L. Carr. Translated by Michael L. Seidler. New York and Oxford: Oxford University Press.
- . 2003. *The Whole Duty of Man According to the Law of Nature*. Edited by Ian Hunter and David Saunders. Translated by Andrew Tooke. Indianapolis: Liberty Fund.
- Smith, Adam. 1976. *The Theory of Moral Sentiments*. Edited by D. D. Raphael and A. L. Macfie. Indianapolis: Liberty Fund.

Bibliography

- Suárez, Francisco. 1679. *Tractatus de Legibus, Ac Deo Legislatore in Decem Libros Distributus, Utriusque Fori Hominibus Non Minus Utilis, Quam Necessarius*. Londini: Sumptibus J. Dunmore, T. Dring, B. Tooke & T. Sawbridge.
- . 2015. *Selections from Three Works*. Edited by Thomas Pink. Translated by Gwladys L. Williams, Ammi Brown, John Waldron, and Henry Davis. Indianapolis: Liberty Fund.

Secondary Sources

- Abramson, Kate. 2015. 'What's So Natural about Hume's Natural Virtues'. In *The Cambridge Companion to Hume's Treatise*, edited by Donald C. Ainslie and Annemarie Butler, 333–68. New York: Cambridge University Press.
- Ainslie, Donald C. 1999. 'Scepticism About Persons in Book II of Hume's Treatise'. *Journal of the History of Philosophy* 37 (3): 469–92.
- . 2007. 'Character Traits and the Humean Approach to Ethics'. In *Moral Psychology*, edited by Sergio Tenenbaum, 94:79–110. Poznań Studies in the Philosophy of the Sciences and the Humanities. Amsterdam and New York: Rodopi.
- . 2015. 'Hume and Moral Motivation'. In *The Cambridge Companion to Hume's Treatise*, edited by Donald C. Ainslie and Annemarie Butler. 283–300. New York: Cambridge University Press.
- Anscombe, G. E. M. 1958. 'Modern Moral Philosophy'. *Philosophy* 33 (124): 1–19.
- Árdal, Páll S. 1966. *Passion and Value in Hume's Treatise*. Edinburgh: Edinburgh University Press.
- Baier, Annette. 1991. *A Progress of Sentiments: Reflections on Hume's Treatise*. Cambridge, Mass and London: Harvard University Press.
- . 2006. 'How Wide Is Hume's Circle?' *Hume Studies* 32 (1): 113–17.
- . 2008. *Death and Character: Further Reflections on Hume*. Cambridge, Mass: Harvard University Press.
- . 2010. *The Cautious Jealous Virtue: Hume on Justice*. Cambridge, Mass: Harvard University Press.

Bibliography

- Bender, Sebastian. 2022. 'Hume's Deep Anti-Contractarianism'. *Hume Studies* 47 (1): 103–29.
- Ben-Moshe, Nir. 2020. 'Hume's General Point of View: A Two-Stage Approach'. *Pacific Philosophical Quarterly* 101 (3): 431–53.
- Boeker, Ruth. 2022. 'Character Development in Shaftesbury's and Hume's Approaches to Self'. In *Hume on the Self and Personal Identity*, edited by Dan O'Brien, 107–30. Cham: Springer International Publishing.
- Brown, Charlotte. 1988. 'Is Hume an Internalist?' *Journal of the History of Philosophy* 26 (1): 69–87.
- . 1994. 'From Spectator to Agent: Hume's Theory of Obligation'. *Hume Studies* 20 (1): 19–35.
- Buckle, Stephen. 1991. *Natural Law and the Theory of Property*. Oxford University Press.
- Buckle, Stephen, and Dario Castiglione. 1991. 'Hume's Critique of the Contract Theory'. *History of Political Thought* 12 (3): 457–80.
- Carlson, Åsa. 2014. 'The Moral Sentiments in Hume's *Treatise*: A Classificatory Problem'. *Hume Studies* 40 (1): 73–94.
- Castiglione, Dario. 1987. 'Hume's Conventionalist Analysis of Justice'. *Annali Della Fondazione Luigi Einaudi* 21:139–73.
- . 1994. 'History, Reason and Experience: Hume's Arguments Against Contract Theories'. In *The Social Contract from Hobbes to Rawls*, edited by David Boucher and Paul Kelly. Abingdon: Taylor & Francis.
- Chamberlain, James. 2017. 'Justice and the Tendency towards Good: The Role of Custom in Hume's Theory of Moral Motivation'. *Hume Studies* 43 (1): 117–37.
- . 2020. 'Hume's Emotivist Theory of Moral Judgements'. *European Journal of Philosophy* 28 (4): 1058–72.
- . 2022. 'Hume on Calm Passions, Moral Sentiments, and the "Common Point of View"'. *Hume Studies* 47 (1): 79–101.
- . forthcoming. 'Hume's General Rules'. Philosophers' Imprint.

Bibliography

- Cohon, Rachel. 2008a. 'Hume's Indirect Passions'. In *A Companion to Hume*, edited by Elizabeth Radcliffe, 159-184. Malden and Oxford: Blackwell Publishing.
- . 2008b. *Hume's Morality: Feeling and Fabrication*. Oxford and New York: Oxford University Press.
- . 2010. 'Hume's Moral Sentiments as Motives'. *Hume Studies* 36 (2): 193–213.
- . 2021. 'Hume's Practice Theory of Promises and Its Dissimilar Descendants'. *Synthese* 199 (1–2): 617–35.
- Darwall, Stephen. 1995. *The British Moralists and the Internal 'Ought': 1640-1740*. Cambridge: Cambridge University Press.
- . 2012. 'Pufendorf on Morality Sociability and Moral Powers'. *Journal of the History of Philosophy* 50 (2): 213–38.
- Edwards, Jeffrey. 2006. 'Hutcheson's "Sentimentalist Deontology"?' *Journal of Scottish Philosophy* 4 (1): 17–36.
- Evnine, Simon. 1993. 'Hume, Conjectural History, and the Uniformity of Human Nature'. *Journal of the History of Philosophy* 31 (4): 589–606.
- Finnis, John. 2011. *Natural Law and Natural Rights*. 2nd ed. Clarendon Law Series. Oxford; New York: Oxford University Press.
- Forbes, Duncan. 1975. *Hume's Philosophical Politics*. Cambridge: Cambridge University Press.
- Frykholm, Erin. 2012. 'The Ontology of Character Traits in Hume'. *Canadian Journal of Philosophy* 42 (S1): 82–97.
- . 2016. 'Associative Virtues and Hume's Narrow Circle'. *Pacific Philosophical Quarterly* 97 (4): 612–37.
- Galvagni, Enrico. 2022. 'Secret Sentiments: Hume on Pride, Decency, and Virtue'. *Hume Studies* 47 (1): 131–55.
- Garrett, Aaron. 2007. 'Francis Hutcheson and the Origin of Animal Rights'. *Journal of the History of Philosophy* 45 (2): 243–65.

Bibliography

- . 2017. ‘Hutcheson on the Unity of Virtue and Right’. In *Kant and the Scottish Enlightenment*, edited by Elizabeth Robinson and Chris W. Surprenant, 19–35. New York and Oxford: Routledge.
- . 2021. ‘Hume, Cicero, and the Ancients’. In *Hume’s An Enquiry Concerning the Principles of Morals: A Critical Guide*, edited by Esther Engels Kroeker and Willem Lemmens, 192–218. Cambridge and New York: Cambridge University Press.
- Garrett, Don. 1997. *Cognition and Commitment in Hume’s Philosophy*. New York and Oxford: Oxford University Press.
- . 2007. ‘The First Motive to Justice: Hume’s Circle Argument Squared’. *Hume Studies* 33 (2): 257–88.
- . 2014. *Hume*. Routledge Philosophers. London and New York: Routledge.
- . forthcoming. ‘Hume’s Geography of Feeling in *A Treatise of Human Nature*’. In *Hume’s A Treatise of Human Nature: A Critical Guide*, edited by Elizabeth Radcliffe. Cambridge: Cambridge University Press.
- Gauthier, David. 1979. ‘David Hume, Contractarian’. *The Philosophical Review* 88 (1): 3–38.
- . 1992. ‘Artificial Virtue and the Sensible Knave’. *Hume Studies* 18 (2): 401–27.
- Gill, Michael B. 2006. *The British Moralists on Human Nature and the Birth of Secular Ethics*. Cambridge and New York: Cambridge University Press.
- . 2014. *Humean Moral Pluralism*. Oxford: Oxford University Press.
- Greco, Lorenzo. 2013. ‘Toward a Humean Virtue Ethics’. In *Aristotelian Ethics in Contemporary Perspective*, edited by Julia Peters, 210–24. New York and Oxford: Routledge.
- Haakonssen, Knud. 1978. ‘Hume’s Obligations’. *Hume Studies* 6 (1): 7–17.
- . 1981. *The Science of a Legislator: The Natural Jurisprudence of David Hume and Adam Smith*. Cambridge and New York: Cambridge University Press.
- . 1996. *Natural Law and Moral Philosophy: From Grotius to the Scottish Enlightenment*. Cambridge: Cambridge University Press.

Bibliography

- . 2000. 'The Character and Obligation of Natural Law According to Richard Cumberland'. In *English Philosophy in the Age of Locke*, edited by M. A. Stewart, 29–48. Oxford: Oxford University Press.
- . 2012. 'Natural Jurisprudence and the Identity of the Scottish Enlightenment'. In *Philosophy and Religion in Enlightenment Britain*, edited by Ruth Savage, 268–78. Oxford: Oxford University Press.
- Harris, James A. 2007. 'Hume's Four Essays on Happiness and Their Place in the Move from Morals to Politics'. *Rivista Di Storia Della Filosofia* 62 (3): 223–35.
- . 2010a. 'Hume on the Moral Obligation to Justice'. *Hume Studies* 36 (1): 25–50.
- . 2010b. 'Introduction: The Place of the Ancients in the Moral Philosophy of the Scottish Enlightenment'. *Journal of Scottish Philosophy* 8 (1): 1–11.
- . 2011. 'Reid and Hume on the Possibility of Character'. In *Character, Self, and Sociability in the Scottish Enlightenment*, edited by Thomas Ahnert and Susan Manning, 31–48. New York: Palgrave Macmillan.
- . 2012. 'The Early Reception of Hume's Theory of Justice'. In *Philosophy and Religion in Enlightenment Britain*, edited by Ruth Savage, 210–30. Oxford: Oxford University Press.
- . 2015. *Hume: An Intellectual Biography*. New York: Cambridge University Press.
- . 2019. 'Hume's Peculiar Definition of Justice'. In *Philosophy, Rights and Natural Law: Essays in Honour of Knud Haakonssen*, edited by Ian Hunter and Richard Whatmore, 216–37. Edinburgh: Edinburgh University Press.
- . 2020. 'Justice in *An Enquiry Concerning the Principles of Morals*'. In *Reading Hume on the Principles of Morals*, edited by Jacqueline Taylor, 77–94. Oxford: Oxford University Press.
- Harrison, Jonathan. 1981. *Hume's Theory of Justice*. Oxford and New York: Oxford University Press.
- Hursthouse, Rosalind. 1999. 'Virtue Ethics and Human Nature'. *Hume Studies* 25 (1–2): 67–82.

Bibliography

- . 2009. 'Hume on Justice'. In *Hume on Motivation and Virtue: New Essays*, edited by Charles R. Pigden, 264–76. New York: Palgrave Macmillan.
- Kopajtic, Lauren. 2015. 'Cultivating Strength of Mind: Hume on the Government of the Passions and Artificial Virtue'. *Hume Studies* 41 (2): 201–29.
- Kroeker, Esther Engels. 2020. 'Hume's *Enquiry Concerning the Principles of Morals* and *The Whole Duty of Man*'. *Journal of Scottish Philosophy* 18 (2): 117–32.
- Latsis, John. 2009. 'Hume and the Concept of Convention'. In *L'invention Philosophique Humienne*, edited by Philippe Saltel, 217–32. Grenoble: Groupe de Recherches Philosophie, Langages & Cognition.
- Mackie, John L. 1980. *Hume's Moral Theory*. London and New York: Routledge.
- Magri, Tito. 1996. 'Natural Obligation and Normative Motivation in Hume's *Treatise*'. *Hume Studies* 22 (2): 231–53.
- McIntyre, Alison. 2014. 'Fruitless Remorses: Hume's Critique of the Penitential Project of *The Whole Duty of Man*'. *Hume Studies* 40 (2): 143–67.
- McIntyre, Jane L. 1990. 'Character: A Humean Account'. *History of Philosophy Quarterly* 7 (2): 193–206.
- . 2006. 'Strength of Mind: Prospects and Problems for a Humean Account'. *Synthese* 152 (3): 393–401.
- Moore, James. 1991. 'The Two Systems of Francis Hutcheson: On the Origins of the Scottish Enlightenment'. In *Studies in the Philosophy of the Scottish Enlightenment*, edited by M. A. Stewart, 37–60. Oxford: Oxford University Press.
- . 1994. 'Hume and Hutcheson'. In *Hume and Hume's Connexions*, edited by M. A. Stewart and John P. Wright, 23–57. Pennsylvania: Pennsylvania State University Press.
- . 2007. 'The Eclectic Stoic, the Mitigated Skeptic'. In *New Essays on David Hume*, edited by Emilio Mazza and Emanuele Ronchetti. Milano: FrancoAngeli.
- Norton, David Fate. 1982. *David Hume: Common Sense Moralist, Sceptical Metaphysician*. Princeton: Princeton University Press.

Bibliography

- . 2005. 'Hume and Hutcheson: The Question of Influence'. In *Oxford Studies in Early Modern Philosophy*, edited by Daniel Garber and Steven Nadler, II:211-256. Oxford: Clarendon Press.
- O'Brien, Dan. 2012. 'Hume and the Intellectual Virtues'. *Discipline Filosofiche* 22 (2): 153–72.
- Pink, Thomas. 2009. 'Natural Law and the Theory of Moral Obligation'. In *Psychology and Philosophy*, edited by Sara Heinämaa and Martina Reuter, 8:97–114. Dordrecht: Springer Netherlands.
- Postema, Gerald J. 1986. *Bentham and the Common Law Tradition*. First edition. Clarendon Law Series. Oxford: Oxford University Press.
- Qi, Xiao. forthcoming. 'Re-evaluating the Principle of Virtuous Motives: Abilities, Justice, and the Concept of Natural Virtue'. *Hume Studies* 50 (1).
- Qu, Hsueh. 2017. 'Hume's Dispositional Account of the Self'. *Australasian Journal of Philosophy* 95 (4): 644–57.
- . 2024. 'Humes Definitions of Virtue'. *Noûs*, 1–19.
- Radcliffe, Elizabeth Schmidt. 1996. 'How Does the Humean Sense of Duty Motivate?' *Journal of the History of Philosophy* 34 (3): 383–407.
- . 2018. *Hume, Passion, and Action*. Oxford: Oxford University Press.
- Russell, Paul. 1995. *Freedom and Moral Sentiment*. Oxford and New York: Oxford University Press.
- Sagar, Paul. 2018. *The Opinion of Mankind: Sociability and the Theory of the State from Hobbes to Smith*. Princeton: Princeton University Press.
- Salter, John. 2015. 'Hume without Spontaneous Order'. *Rationality, Market and Morals* 6:26–38.
- Santos Castro, Juan Samuel. 2017. 'Hume and Conjectural History'. *Journal of Scottish Philosophy* 15 (2): 157–74.
- Sayre-McCord, Geoffrey. 2016. 'Hume on the Artificial Virtues'. In *The Oxford Handbook of Hume*, edited by Paul Russell, 435–69. New York: Oxford University Press.

Bibliography

- Schneewind, J. B. 1990. 'The Misfortunes of Virtue'. *Ethics* 101 (1): 42–63.
- . 1998. *The Invention of Autonomy: A History of Modern Moral Philosophy*. Cambridge and New York: Cambridge University Press.
- Shaver, Robert. 1992. 'Hume on the Duties of Humanity'. *Journal of the History of Philosophy* 30 (4): 545–56.
- Sugden, Robert. 2005. *The Economics of Rights, Cooperation and Welfare*. Hampshire and New York: Palgrave Macmillan.
- . 2009. 'Can a Humean Be a Contractarian?' *Rationality, Market and Morals* 0:11–23.
- Swanton, Christine. 2015. *The Virtue Ethics of Hume and Nietzsche*. West Sussex: Wiley-Blackwell.
- Taylor, Jacqueline. 1998. 'Justice and the Foundations of Social Morality in Hume's *Treatise*'. *Hume Studies* 24 (1): 5–30.
- Walschots, Michael. 2022. 'Hutcheson's Theory of Obligation'. *Journal of Scottish Philosophy* 20 (2): 121–42.
- Wand, Bernard. 1956. 'Hume's Account of Obligation'. *The Philosophical Quarterly* 6 (23): 155–68.
- Watkins Tate, Margaret. 2005. 'Obligation, Justice, and the Will in Hume's Moral Philosophy'. *Hume Studies* 31 (1): 93–122.
- Westerman, Pauline. 1994. 'Hume and the Natural Lawyers: A Change of Landscape'. In *Hume and Hume's Connexions*, edited by M. A. Stewart and John P. Wright, 83–104. Pennsylvania: The Pennsylvania State University Press.
- Whelan, Frederick G. 1994. 'Hume and Contractarianism'. *Polity* 27 (2): 201–24.
- Wiggins, David. 1991. 'Categorical Requirements: Kant and Hume on the Idea of Duty'. *The Monist* 74 (1): 83–106.
- . 1998. 'Natural and Artificial Virtues'. In *How Should One Live?*, edited by Roger Crisp, 131–40. Oxford: Oxford University Press.
- Williams, Bernard. 2006. *Ethics and the Limits of Philosophy*. London: Routledge.