THE PSYCHOLOGY OF LOYALTY: A SOLUTION TO HUME’S
PROBLEM OF MORAL OBLIGATION

by

AVITAL HAZONY

Copyright © Avital Hazony 2022

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PHILOSOPHY

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2022
THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation
prepared by: Avital Hazony
 titled: The Psychology of Loyalty: A Solution to Hume’s Problem of Moral Obligation

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of
Doctor of Philosophy.

Michael B. Gill  
Michael B. Gill  Date: Aug 3, 2022

Mark Timmons

Mark Timmons  Date: Aug 2, 2022

Sara Aronowitz
Houston Smit

Date: Aug 4, 2022

Houston Smit
Date: Aug 3, 2022

Final approval and acceptance of this dissertation is contingent upon the candidate’s submission
of the final copies of the dissertation to the Graduate College.

We hereby certify that we have read this dissertation prepared under our direction and
recommend that it be accepted as fulfilling the dissertation requirement.

Michael B. Gill  
Michael B. Gill  Date: Aug 3, 2022

University of Edinburgh

Mark Timmons

Date: Aug 2, 2022

Philosophy
Acknowledgements

I would like to express my gratitude to those who assisted me in the journey of writing this dissertation. My greatest debt is to my advisor, Michael Gill, who has read countless drafts and discussed many of my ideas with me for the past five years. His optimism and curiosity are inspiring, and in his comments he has always challenged me to be true to human nature. I am also indebted to Mark Timmons, co-chair of my committee, who greatly shaped this dissertation by introducing me to W.D. Ross and suggesting that I write a chapter on loyalty. I want to thank Eric Schliesser for his encouragement and for often saving me from being blinded by my own theory. I am grateful to Sara Aronowitz for taking on my project without ever having met me, and for giving invaluable insight into how to clarify my terms and improve my writing. I am grateful to Houston Smit, and to all the professors on my committee, for devoting time and energy to help me improve as a philosopher. I also want to thank the editors and three anonymous reviewers at *Hume Studies* for their comments on earlier versions of chapter 1, and Hume scholars Lorraine Besser and Jackie Taylor for their comments on earlier versions of chapter 2.

This dissertation is a product of many years of philosophical learning. In this I am indebted first to my father, Yoram Hazony, who encouraged my questioning, introduced me to philosophy, and is still a role model for how to offer new answers to ancient questions. I am grateful to Joshua Weinstein, who helped me think through my very first proposal on moral obligation. I want to thank the professors who encouraged and guided me in the early stages of my career in academic philosophy: Andrew German, Gideon Rosen, Jeffrey Stout, and Margaret Watkins. I am grateful to all the professors and graduate students at University of Arizona who
offered me an intellectual home. I especially want to thank David Owens, who found time to study Book I of the *Treatise* with me, and Julia Annas, who gave me valuable feedback on my writing. Mario Ivan Juarez Garcia, with whom I argued about my ideas, has become a true friend, and teaches me how to navigate the web of relationships that is academic philosophy. Finally, I am greatly indebted to Sandra Kimball, Academic Program Manager at the University of Arizona’s philosophy department, for her constant guidance.

The centrality of loyalty to morality was not an idea I came up with on my own. I believe it came from growing up in the house of my parents, Yoram and Yael Hazony, who were exemplars of loyalty in their marriage and in their upholding of the Jewish tradition. They also gave me the incredible gift of having eight younger brothers and sisters, thereby teaching me early on the meaning of the extended self. Through conversation and expectation, they taught me to view myself as empowered to act on behalf of others who are part of my extended self and to view philosophical enquiry as directed towards being a better person. I am also thankful to my siblings, who argue with me and are loyal in return. Finally, I am grateful to God for bringing my husband Eliyahu Levi into my life. Eliyahu encourages me when I most need support. Moreover, he is a true philosopher, willing to fight for truth when he finds it, and in that he offers me companionship and confidence to continue in my work.
Table of Contents

ABSTRACT .................................................................................................................................................. 7
INTRODUCTION............................................................................................................................................. 8

Chapter 1: Hume’s Theory of Moral Judgment in Light of His Explanatory Project ........... 16
  1. Two Kinds of Moral Approval in Hume’s Account of Justice ...................................................... 17
  2. The Unity of Hume’s Explanation of Moral Judgment ................................................................. 26
  3. Contribution of Non-Approved Motives to Constitution of the Artificial Virtues ............ 36

Chapter 2: Humean Loyalty ......................................................................................................................... 48
  1. Hume’s Extended Self ....................................................................................................................... 49
  2. Humean Loyalty ............................................................................................................................... 57
  3. Distinguishing Humean Loyalty from Humean Sympathy .......................................................... 66
  4. Critique of Hume’s Explanation of Love of Relations ................................................................. 70
  5. The Virtue of Loyalty ....................................................................................................................... 75
  6. Objection: Hume’s Disapproval of Faction ................................................................................... 80

Chapter 3: Loyalty as a Solution to Hume’s Problem of Moral Obligation ...................... 86
  1. Critique of Hume’s Judgment Internalism ...................................................................................... 89
     1.1 Hume’s Problem of Obligation ................................................................................................. 90
     1.2 Hume’s Judgment-Internalist Solution to the Problem of Obligation .............................. 95
     1.3 A Critique of Hume’s Solution ............................................................................................... 99
     1.4 A Motive for Obligation .......................................................................................................... 106
  2. Loyalty: The Impression that Gives Rise to the Idea of Obligation .................................... 110
     2.1 Hume’s Experimental Method ............................................................................................... 111
     2.2 Hume’s Method Stage A: Describing the Phenomena of Obligation ............................. 113
     2.3 Hume’s Method Stage B: Identifying Loyalty as the Common Impression ............... 116
     2.4 Loyalty Motivates a Subset of Moral Actions ........................................................................ 119
     2.5 The Strength of Loyalty ........................................................................................................ 122
2.6 Hume’s Method Stage C: Causes of Loyalty and Explanation of Obligation …… 123

3. Rethinking Humean Moral Judgment ............................................................. 127
   3.1 Unity and Pluralism in the Humean Idea of Obligation ............................... 128
   3.2 The Function of Obligation .......................................................................... 131
   3.3 Calibrating the Two Axes of Moral Judgment .............................................. 138
       3.3.1 The Two Axes .......................................................................................... 139
       3.3.2 Calibration in Judgment of Action .............................................................. 141
       3.3.3 Calibration in Extending the Self ............................................................... 143
       3.3.4 Calibration in Evaluating Loyalties .......................................................... 145

Chapter 4: Group Loyalty and the Psychology of Impersonal Obligations .......... 152
   1. Recognizing Group Loyalty by Its Causes ..................................................... 153
   2. Challenging the Impartiality of Impersonal Obligations ............................... 158
       2.1 Hume on Limiting Partiality ..................................................................... 159
       2.2 Partiality in the Adoption and Moral Approval of the Convention ............ 161
           2.2.1 Partiality in Adopting the Convention ................................................... 164
           2.2.2 Partiality as Prerequisite to Moral Approval of Convention ................. 169
       2.3 Rejection of the Interpretation that Conventional Obligation Comes From Rules. 173
   3. Group Loyalty as the Natural Motive to Act According to the Convention ........ 179
       3.1 Common Interest and Trust ..................................................................... 181
       3.2 Doing One’s Share in A Common Enterprise ............................................. 184
       3.3 Loyalty to the Group that Shares the Convention ....................................... 186
       3.4 A Comparative Note .................................................................................. 190
   4. The Psychology of Impersonal Obligations ................................................... 192
   5. Preliminary Thoughts on Normative Implications .......................................... 199

CONCLUSION ........................................................................................................ 208

BIBLIOGRAPHY .................................................................................................. 213
ABSTRACT

This dissertation argues that moral obligations are derived from loyalty. The argument is framed as an answer to David Hume’s psychological problem of moral obligation: what is the impression from which the idea of obligation is derived. The proposal is that the idea of obligation is derived from the reflective impression of loyalty, and therefore that moral obligations are acts that fulfill the needs of those who are part of the agent’s extended self.

Chapter 1 argues that Hume’s theory of moral judgment is best understood in light of his explanatory method. It shows that Hume aims to explain moral judgment by distinguishing the impression that gives rise to the concept of virtue and finding the causes of this impression.

Chapter 2 develops a Humean theory of loyalty. Hume’s discussion of pride suggests a view of the self as extending to include others, and this view of the self is used to define loyalty as the tendency to act on the motive to fulfill the needs of those who are part of the agent’s extended self. Chapter 3 critiques and supplements Hume’s view of moral obligation. It argues that Hume’s moral sentiments cannot explain obligation, on their own, and that loyalty is the impression that gives rise to the idea of obligation. This leads to a new theory of moral judgment of obligation as a calibration of agent-neutral moral sentiments and agent-relative loyalty. On this view moral obligations are always acts that benefit and fulfill the needs of those who are part of the agent’s extended self. Chapter 4 responds to the challenge that loyalty cannot explain impersonal obligations. Analysis of Hume’s account of justice shows that on Hume’s view moral rules are only obligatory if they are developed and maintained as conventions of a particular group. Loyalty to the group who shares the convention defines impersonal obligations.
Hume’s charge against all moral theories before him was that they didn’t explain moral obligation. He thought that previous moral theories explained where our distinctions of virtue and vice come from, but did not explain why they obligate. Hume argued that the explanation of moral obligation, or the binding nature of morality, does not just follow from explaining what is good and what is bad, but needs to be explained separately: “as this ought or, ought not, expresses some new relation or affirmation, `tis necessary it shou’d be observ’d and explain’d.”

Hume’s charge at first seems to be a curious one, since many moral theories attempt to show that morality is not merely a preference, but rather binding. But Hume is not asking what could convince people to act morally, or why acting morally is justified. What is unique about Hume’s question is that it is a psychological question. Hume thought we only explain why morality is obligatory if we explain how it is connected to our will: “to prove that the measures of right and wrong are eternal laws, obligatory on every rational mind… We must also point out the connexion between the relation and the will.” Hume thought that we would explain moral obligation only if we pointed to the motive that connects our moral judgment with our will to act morally. I call this Hume’s problem of obligation.

---

1 T 3.1.1.27, 469-70. All references to the Treatise are to David Hume, A Treatise of Human Nature, L.A. Selby-Bigge (ed.), Oxford University Press, 1978. References to the Treatise will be marked with a T, followed by the section numbers and page number of this edition. References to An Enquiry Concerning the Principles of Morals will be marked with an M, followed by the section and page number of the Selby-Bigge edition. Both are available online at davidhume.org/texts/. References to Hume’s other essays will be to David Hume, Essays Moral, Political and Literary, Liberty Fund, Indianapolis, 1985.
2 T 3.1.1.22, 465-6.
Hume suggested a solution to the problem of obligation when he argued that the ideas of virtue and vice employed in moral judgment have their source in moral sentiments, which themselves motivate. Thus his answer to the problem of obligation was that obligation is explained by the moral sentiments. In this dissertation I argue that Hume’s solution to the problem of obligation was inadequate, because he did not study the phenomena of moral obligation separately from virtue and vice. I show that in order to understand the connection between our moral judgments and the will, we must study instances in which acting morally is hard. In these instances, I contend, we find that the moral sentiments do not sufficiently motivate, but rather are strengthened by the motive of loyalty. Thus moral sentiments on their own do not explain the connection between our moral judgments and our will to act morally. The solution to Hume’s problem of obligation is that loyalty influences our will and gives rise\(^3\) to the idea of obligation. Thus loyalty is the psychological source of obligation. Moral obligation is explained by the calibration of loyalty with the moral sentiments.

In order to argue that obligations are explained by loyalty, I offer a novel theory of loyalty based in Hume’s psychology. I first show that the self, on Hume’s psychology, extends to include other people and groups. Those who are part of the extended self are those we consider ours – our family, our friends, our university, our team, our country. This extension of the self to include others offers a motive for action that is different from either self-interest or altruism. I define loyalty in the following way:

\(^3\) I use the terminology “giving rise to” following Hume’s own use when referring to the process by which we acquire ideas from impressions (see for example T 1.3.2.4, 74-5). Hume says that ideas are either copied from impressions or represent impressions. Simple ideas are copied from single impressions, but abstract ideas represent a group of simple ideas. See T 1.1.7.1-1.1.7.2, 17-18; Don Garrett, *Cognition and Commitment in Hume’s Philosophy*, Oxford University Press, Oxford, 1997, 23-25.
Loyalty: acting on behalf of the needs and interests of those who are part of the agent’s extended self, when motivated by the fact that they are part of the agent’s extended self.

I then use this theory of loyalty to argue that when we feel that we ought to do something, or that we have to do it even though we don’t want to, what we are feeling is the motive of loyalty. By demonstrating this role of loyalty in examples of obligation, I show that the impression that gives rise to the idea of obligation is the motive of loyalty. This conclusion leads me to rethink Hume’s account of moral judgment. While Hume held that moral judgment has its source in moral sentiments, I argue that the judgment that we have a moral obligation which is binding is a calibration of moral sentiments with loyalty.

This study of Humean obligation combines an interpretation of Hume’s moral theory in the *Treatise*, a critique of Hume’s account of obligation, and an application of Hume’s experimental method to studying the question of moral obligation. In this way I hope to take the best aspects of Hume’s moral theory and method, while utilizing them to answer a question that Hume raised but did not adequately answer. In order to understand Hume’s experimental method, the reader should have in mind the basics of Hume’s psychology. I use the terms perception, idea, impression, and reflective impression as Hume used them, using the following division:⁴

---

⁴ T 1.1.1.1, 1-2.
In the dissertation I use the terms idea and concept interchangeably. I hope that translating Hume’s term to the modern one is helpful rather than confusing to the reader. An empiricist about ideas, Hume argued that all ideas (thoughts) are derived from impressions (feelings). Impressions, which are the source of ideas, include both impressions of sense, which are not preceded by other perceptions, and reflective impressions, which are preceded by another impression or idea. Thus pains and pleasures, colors and tastes, which are impressions of sense, are followed by desires, passions and sentiments, which are reflective impressions. Reflective impressions influence our will and motivate action.

In chapter 1 I discuss, and in chapter 3 I utilize, Hume’s experimental method. This method explains ideas we use in thought by pointing out the impressions from which the ideas are copied and then finding the causes of these impressions. Hume’s view is that by tracing from the ideas used in thought to the impressions that give rise to them, and considering these impressions in the contexts in which they appear, we are able to offer a causal explanation of our ideas. Crucially, Hume argued that moral ideas have their source not in impressions of sense (like the idea of red) but rather in reflective impressions (like the idea of pride). My application of Hume’s experimental method to the idea of obligation will thus focus on ascertaining the reflective impression that gives rise to the idea of obligation and finding the causes of this reflective idea.

Hume’s method is based on his distinction between ideas and impressions, which are not used in psychology and cognitive science today. This may lead to the worry that his method
is outdated.⁷ I think that Hume’s experimental method can be useful even if we have a different view of the mind. As Hume says, “Every one of himself will readily perceive the difference betwixt feeling and thinking.”⁸ Hume’s method is useful because it offers a causal explanation of ideas that we use in thought by tracing to a more fundamental source and placing it in the context in which it occurs. By doing so, Hume holds both that our ideas are not fundamental or innate, and that they have a genealogy which forms them. The method of tracing from ideas to their impressions allows us to become aware of the causal genealogy of our ideas, while at the same time not giving too much importance to the causes in and of themselves. The causes are only important because they are the source of the ideas we find useful in thought. This method allows us to be skeptical about the way we use ideas without assuming that we can replace them at will.

Having said that, the conclusions I arrive at – that moral obligation arises from loyalty, and that the judgment of moral obligation is a calibration of moral approval and loyalty – do not depend on Hume’s method. I propose these as an amendment to Hume’s sentimentalist moral psychology, but they offer an independent theory of the psychology of obligation. My conclusion has import for both meta-ethics and normative ethics. In meta-ethics, I argue for a judgment internalist view of the judgment of moral obligation. On this view moral judgment of an action is not achieved by weighing countless reasons for action. Rather it has a certain structure, bringing our moral approvals to bear on actions that fulfill the needs of our extended self, and thus achieves the prioritization needed to produce action. In normative ethics, my view

---

⁷ For example, for a long time it was thought that Hume thought perceptions are pictorial, a view that has been discarded in cognitive science. Stephanie Rocknak has argued that this was not Hume’s view. Stephanie Rocknak, *Imagined Causes: Hume’s Conception of Objects*, Springer, New York, 2013, 15-26.

⁸ T 1.1.1.1, 1-2.
entails that since obligations are explained by the needs of those who are part of our extended self, obligations are not universal.

Chapter 1 aims to better understand Hume’s theory of moral in light of his explanatory project. It argues that both the virtue theory reading and the natural law reading unduly narrow Hume’s theory of moral judgment as it is presented in his account of justice. I show that Hume thought we morally evaluate both individuals’ characters, by evaluating their motives, and kinds of actions, by evaluating them in context of systems of conduct. I then argue that Hume’s pluralist account of moral judgment can be unified if we view it in light of Hume’s explanatory project. Hume argues that moral judgment utilizes ideas of virtue and vice that have their source in impressions of moral sentiments, which are in turn caused by agreeableness and utility to self and others. These causes, I argue, characterize both motives of individuals and systems of conduct, and therefore offer a unified explanation of moral judgment.

In chapter 2 I present a novel theory of loyalty, which offers the psychological basis for my argument that loyalty is the impression that gives rise to the idea of obligation. By analyzing Hume’s account of the self, as it is found in his account of pride, I argue that the Humean self does not have set boundaries and extends to include other people, groups and institutions. The needs and interests of these people who are part of the extended self are the causes of a particular reflective impression which is the motive of loyalty. I argue that loyalty is not strong Humean sympathy, since sympathy is caused by an association of a sentiment of another person with the self, while loyalty is caused by the association of another person with the self. This difference in association leads to the different roles sympathy and loyalty play in our moral lives. While Humean sympathy is the mechanism by which we come to have Humean
moral sentiments, and thus offers the basis for Hume’s account of the judgment of virtue and vice, loyalty is a central motive to act virtuously. Finally, I argue that loyalty is a virtue on Hume’s view of the virtues, since it is useful and agreeable to both self and others.

In chapter 3 I turn to argue that loyalty is the psychological source of obligation. I begin chapter 3 by explaining Hume’s problem of obligation and his solution to it. Hume offers a judgment internalist view of obligation, on which moral sentiments are the reflective impression that give rise to the ideas of virtue and vice, and he assumes that this account also explains obligation. I critique Hume’s solution by arguing that we can empirically distinguish between moral approval and obligation. I follow W.D. Ross in proposing that we must distinguish two moral concepts, the good and the obligatory. I then apply Hume’s experimental method to the idea of obligation. I first investigate instances of moral obligation, and show that they share a common motive of loyalty that motivates when moral approval is too weak to motivate action. By tracing to the causes of loyalty I conclude that obligations are actions that fulfill the needs of those who are part of the extended self. I then argue that in order to maintain Hume’s judgment internalism, we must rethink Hume’s account of moral judgment. I distinguish between moral evaluation, which has its source in moral sentiments alone, and moral judgment, which calibrates the moral sentiments with loyalty. I argue that our moral judgment, which motivates moral action, is a judgment that combines moral evaluation which is agent-neutral with loyalty which is agent-relative. The function of moral judgment is to prioritize among moral evaluations, by calibrating them with the associations that make up the agent’s self, in order to successfully motivate moral action.
Having argued that obligation has its psychological source in loyalty, I turn to a central objection to this view: that loyalty cannot explain impersonal obligations. In chapter 4 I argue that it is the psychology of group loyalty, and not the ability to apply rules, which explains impersonal obligations. I first argue that we can be loyal to groups, while we cannot be loyal to rules, because the causes of loyalty are the needs of those who are part of the agent’s extended self. I then turn to Hume’s account of justice and show that Hume thinks the impersonal obligations pertaining to property arise only in context of a convention. This convention, I argue, is a system of conduct adopted by a particular group of people. This reading of Hume presents him as a non-Kantian, who holds that rules do not obligate in and of themselves but rather by virtue of their adoption by, and utility for, a certain group or society. I then use this account Hume gives of impersonal obligation to argue that personal obligations and impersonal obligations have the same psychological source, loyalty, although the former arise from loyalty to individuals and the latter arise from loyalty to the group who shares the convention. Finally, I offer some preliminary thoughts on the implication of my view for normative ethics. What follows from my explanation of obligation is that we can only ascribe to people obligations to act on the needs of those who are part of their extended self. This means that obligations are not universal, but rather depend on the agent’s extended self.
Chapter 1: Hume’s Theory of Moral Judgment in Light of His Explanatory Project

Two different interpretations of Hume’s theory of moral judgment are often presented in discussions of Hume’s account of justice in *A Treatise on Human Nature*. The virtue interpretation holds that, according to Hume, justice is a virtue because we approve of a certain motive to act justly. On the natural law interpretation, Hume thinks we morally approve of justice because we approve of a system of actions, regardless of their motives. In this chapter I show that both of these interpretations capture aspects of Hume’s account of moral judgment in the case of justice, and that combining them will give a full picture of Hume’s theory of moral judgment. I argue that Hume’s theory of moral judgment is pluralistic, including judgments of both character and systems of action, but it can be unified by the causes of the moral sentiments. This conclusion clarifies Hume’s theory of moral judgment and shows that it is best understood in light of his explanatory project.

In part 1 I lay out the textual support to show that Hume’s account of justice includes both approval of a motive and approval of a system of actions. In part 2 I argue that we can allow such plurality in Hume’s theory of moral judgment if we view Hume’s account of moral judgment in light of his explanatory project. Hume offers a unified theory of moral judgment because he can show that these different approvals are explained by the same causes. Finally, in part 3 I argue that viewing Hume’s account of moral judgment in light of his

---

9 I will focus on Hume’s account of justice in the Treatise, since it has a more extensive discussion of motivations. At times I refer to the account in *An Enquiry Concerning the Principles of Morals*, which I do not think is very different on the main points of my argument. But I will avoid making claims about the Enquiry, since some hold that Hume changes his moral theory dramatically in that work. For this view see Jacqueline Taylor, “Justice, Sympathy and the Command of our Esteem”, *Diametros* 44, 2015, 173-188.
explanatory project allows us to appreciate a further distinction between the natural and the artificial virtues: while judgements in the case of the former are fully explained by the approval of a certain motive, judgments in the case of the latter are only fully explained by approval of a motive in the context of certain non-approved motives that contribute to the constitution of a convention.

1. Two Kinds of Moral Approval in Hume’s Account of Justice

In this part I present two different interpretations of the role of approval of motives in Hume’s account of justice: the virtue theory reading and the natural law reading. I argue that we must combine these approaches, in order to get a full picture of Hume’s account of moral judgment in the case of justice. The textual evidence leads me to conclude that Hume thinks there is both a virtue of justice, which arises from approval of a particular motive, and approval of acting in accordance with a convention of justice, regardless of motive, because it is beneficial to society. I therefore suggest that Hume gives a pluralist account of moral judgment in the case of justice: we morally approve of both a particular motive and of a system of conduct. On this reading of Hume, approval of motives only explains some of our moral judgments.

Hume’s account of virtue is sentimentalist: it is approval or disapproval of a person’s motive from a general and common point of view which is the source of the ideas of virtue and vice employed in moral judgment. This approval or disapproval is a sentimental reaction to a particular person’s motivation, not to a disembodied moral principle. The motives we approve of, such as beneficence or courage, are not a motives to be virtuous, meaning they are motives that require the employment of concepts of virtue or vice. Rather it is the approval of
these motives that renders the person who is motivated by them virtuous and gives rise to the concept of virtue.\textsuperscript{10}

Hume considers it evidence for his view that we don’t primarily judge actions virtuous or vicious. We judge people virtuous or vicious by approving or disapproving of their motives. Actions are judged virtuous or vicious derivatively, depending on the approval or disapproval of the motive that brings them about: “all virtuous actions derive their merit only from virtuous motives, and are consider’d merely as signs of those motives.”\textsuperscript{11} Hume brings examples to substantiate this claim: when we morally approve of a person’s motive, we judge him virtuous even if he is not able to carry out the action he is motivated to pursue. Additionally, if a person is motivated by a malicious motive to act in a way that appears virtuous, we will not consider the person virtuous.\textsuperscript{12} Thus Hume argues that virtues are defined by their motives, and moral judgment – to the extent that it is a judgment of vice or virtue – has motives as its object.

The role of approval of motives seems straightforward in this account. But when Hume comes to his account of justice, he distinguishes between two questions regarding motives: he distinguishes between the genetic question, “the manner, in which the rules of justice are establish’d by the artifice of man”, and the moral question, what are “the reasons, which determine us to attribute to the observance or neglect of these rules a moral beauty and deformity.”\textsuperscript{13} Hume separates the genetic question, where he asks what motivation brings about the institution of justice, and the moral judgment question, where he asks what motivation we

\textsuperscript{10} For a helpful discussion of this point see Sayre-McCord, “Hume on the Artificial Virtues”, 452-453.
\textsuperscript{11} T 3.2.1.4, 478. Although actions are judged on the merit of their motive, we also take actions to be signs that help us judge the motivations of a person, since we cannot “look within to find the moral quality.” T 3.2.1.2, 477.
\textsuperscript{12} T 3.2.1.3, 477-478.
\textsuperscript{13} T 3.2.2.1, 484.
approve of that makes a person just. Notice this distinction has led to the elucidation of the difference between Hume’s genetic account, on which self-interest motivates people to seek cooperation and adopt a convention of clear property boundaries, and his moral account, on which we need a motive human beings approve of to explain the judgment of virtue.

While Hume says that self-interest motivates people to institute the convention of justice, he never clearly names the motive of the just person of which we approve. This local interpretive problem leads to a more general debate: does Hume’s discussion of justice support his virtue theory, on which moral judgment has its source in approval of motives? If Hume points to, or can be read as accepting, an approved motive that renders justice a virtue, then this would support the view that Hume’s moral theory is primarily a theory of virtues and that moral judgment is defined by approved motives. But if Hume doesn’t point to an approved motive of justice, then it seems that justice, and perhaps the other artificial virtues as well, are a departure from Hume’s theory of the virtues. Since the artificial virtues are a major part of Hume’s moral theory, if they do not fit into Hume’s theory of virtues this implies that Hume is not committed to

---

15 Hume does not think self-interest is ever limited to each individual caring only about him or herself. Human self interest also motivates to act on behalf of family and friends. See T 3.2.2.5, 486-487.
16 See T 3.2.2.24, 498-500 and T 3.2.8.5, 543-544.
17 This question is raised by Christine Swanton, “Hume and the Problem of Justice as a Virtue”, in *The Virtue Ethics of Hume and Nietzsche*, Wiley Blackwell, Chichester, 2015, 71.
18 Swanton distinguishes between a virtue ethicist, who thinks that we judge the worth of actions solely on the basis of what a virtuous person would do, and a philosopher who puts forth a theory of virtue. See Christine Swanton, “Can Hume be Read as a Virtue Ethicist?”, *Hume Studies* 33(1), 2007, 93.
19 Some hold that for Hume to be a virtue ethicist he must think that the virtues are required for human flourishing and that the virtues do not conflict with one another, as argued by Rosalind Hursthouse, “Virtue Ethics and Human Nature”, *Hume Studies* 25(1-2), 1999, 67-82. Both of these claims are difficult to attribute to Hume, as argued by Harris, “Hume on the Moral Obligation to Justice”, 30-31. Since the virtue ethics reading is not primarily focused on motives, it less important to my discussion here. My interest is in the role motives have in construing justice a virtue, and on the implication this has for Hume’s moral theory, and this is why I focus on Garrett and Sayre-McCord’s readings that are not explicitly committed to a virtue ethics reading of Hume.
a virtue theory, and we may have to seek another unifying source for his theory of moral judgment in the Treatise.\(^{20}\)

Don Garrett and Geoffrey Sayre-McCord argue that justice fits into Hume’s virtue theory because justice, like the natural virtues, has a non-moral motive (a motive that has no specifically moral content)\(^{21}\) of which we approve and which renders justice virtuous. This view may at first seem to conflict with Hume’s claim that “we have naturally no real or universal motive for observing the laws of equity, but the very equity and merit of that observance.”\(^{22}\) But Garrett points out that the word “naturally” was added by Hume in later editions, and suggests that Hume means to emphasize that there is no natural motive, rather than no motive at all, that we approve of.\(^{23}\) Garrett and Sayre-McCord therefore take Hume to mean that there is no natural motive for justice before the institution of the convention of justice. Rather, the motive of justice of which we approve is developed in response to an existing convention of private property. As Hume says, “the conventions of men… create a new motive.”\(^{24}\) This new motive that develops in

\(^{20}\) There is another possibility, that Hume is an error theorist, as suggested by J.L. Mackie, *Hume’s Moral Theory*, Routledge and Kegan Paul Ltd, Norwich, 1980. Marcia Baron continues this line of interpretation when she argues that justice is artificial because it involves a lie. People must believe that respecting property is good, in order to bring about its beneficial consequences, but in fact they are motivated by education and political power, not by any approved motive, and therefore justice is not truly a virtue. See Baron, “Hume's Noble Lie: An Account of His Artificial Virtues”, *Canadian Journal of Philosophy*, 12 (3), 1982, 539-555. I agree with Baron that not all just actions are motivated by an approved motive, but disagree with her that this precludes the possibility of there being a virtue of justice that is constituted by an approved motive.

\(^{21}\) For discussion see Sayre-McCord, “Hume and the Artificial Virtues”, 439-441.

\(^{22}\) T 3.2.1.17, 483.

\(^{23}\) Garrett, “The First Motive to Justice”, 260. Sayre-McCord concurs, “Hume and the Artificial Virtues”, 444. See Treatise 2.2.2.8, 335: “in vain should we expect to find, in *uncultivated nature*, a remedy for this inconvenience; or hope for any inartificial principle of the human mind, which might control those partial affections… The idea of nature can never serve this purpose, or be taken for a natural principle, capable of inspiring men with an equitable conduct towards each other”. See also T 2.2.2.21, 343: “those impressions, which give rise to this sense of justice, are not natural to the mind of man, but arise from artifice and human conventions.”

\(^{24}\) T 3.2.5.10, 521-522.
response to the existence of a convention is an artificial motive, and it is this motive we approve of.25

Garrett thus reads Hume as concluding that the motivation to be just, in a society that has conventions of property, is the desire to regulate one’s actions by the rules of justice and exclude other motives for action. This is how Garrett understands Hume’s reference to a motive of “regard to justice.”26 This desire to regulate one’s actions by the rules of justice is an artificial motive because it only appears in human psychology in response to the existence of a convention, and does not exist in human psychology in a society that has no conventions.27 Thus Garrett maintains that Hume sticks to his virtue theory, since justice fits into his theory that virtues are constituted by an approved motive.28 Justice fits into Hume’s account of the virtues because justice is constituted by an approval of a motive of the person who is judged to be just. Justice is an artificial virtue because the motive we approve of is an artificial one.

Sayre-McCord agrees with Garrett that Hume sticks to his theory of virtue.29 But Sayre-McCord disagrees with Garrett on what the approved artificial motive is. Sayre-McCord objects to Garrett’s suggestion that the motive is regard for the rules, since he thinks the motive of rule following can motivate following rules that are not mutually advantageous. This, he argues, doesn’t fit Hume’s view that we only approve of following conventions which are

26 T 3.2.1.9, 479-480 and T 3.2.1.11, 480-481.
mutually advantageous. Sayre-McCord therefore offers Hume a motive that Hume himself does not discuss, the motivation to do one’s share in a cooperative scheme. This is the motive we approve of which constitutes the virtue of justice. Although Sayre-McCord disagrees with Garrett on the specific approved motive, he agrees with Garrett that justice fits into Hume’s virtue theory framework, because there is an artificial motive of which we approve that renders justice a virtue.

James Harris opposes the interpretation offered by Garrett and Sayre-McCord, which tries to fit justice into Hume’s account of the virtues. Harris emphasizes that Hume never specifies a motivation we approve of for justice, and therefore concludes that Hume is not concerned with fitting justice into a virtue theory. Harris argues instead that Hume is writing in the natural law tradition, which is concerned with rights and obligations, and follows Hugo Grotius who says explicitly that motivations are irrelevant when it comes to judging acts of justice. Moreover, Harris thinks Hume’s use of virtue terminology does not reflect a commitment to a moral theory of virtue.

Harris’ view at first seems problematic, since Hume reiterates his virtue theory in the opening passages of his chapter on justice. Rachel Cohon therefore attempts a middle ground,

31 This is also the path Christine Swanton takes when she suggests that the natural motive of justice is compassion. See Swanton, “Hume and the Problem of Justice as a Virtue”, 78-86.
32 Sayre-McCord distinguishes doing one’s share and doing one’s fair share, in order to maintain Hume’s distinction between the approved, non-moral motive as it appears before the existence of a convention and the moral, approved motive as it appears once the conventions of justice are in place. See “Hume on the Artificial Virtues”, 452.
33 Ibid, 452.
35 Harris, “Hume on the Moral Obligation to Justice”, 27-28. “Hume’s view is that in the case of justice it does not matter what the motive is, just so long as it issues in a reliable disposition to abide by the conventions of justice.”
arguing that Hume maintains a moral theory of virtue although he does not point to an approved motive. She does this by suggesting that Hume thinks it is not approval of a motive that constitutes a virtue, but approval of a character trait. Cohon agrees with Garrett and Sayre-McCord that Hume sticks to a virtue theory, but she thinks Hume mentions that virtue is constituted by an approved motive because it was a commonly held view, not because it is his own. Cohon argues that Hume’s view is that virtues are constituted by approved character traits. This allows Cohon to read Hume as maintaining his virtue theory, without glossing over the fact that he does not specify an approved motive for justice.\(^\text{37}\)

Margaret Watkins has challenged Cohon’s claim that Hume thinks virtues arise from approval of a trait, and not a motive.\(^\text{38}\) But what is more important to our discussion here is that approval of a trait or a motive cannot fully explain moral judgment in the case of justice, on Hume’s account of justice, because Hume explicitly relies on approval of consequences in his account of justice. As Harris points out, Hume thinks we approve of acting according to the convention of justice because we approve of the consequences of the convention. Harris emphasizes the following paragraph, in which Hume makes this claim:

Tho’ justice be artificial, the sense of its morality is natural. ’Tis the combination of men, in a system of conduct, which renders any act of justice beneficial to society. But when once it has this tendency, we naturally approve of it; and if we did not so, ’tis impossible any combination or convention cou’d ever produce that sentiment.\(^\text{39}\)

\(^{38}\) Margaret Watkins Tate makes a strong case against this reading when she points out that Hume discusses the motivation for virtue in 2.2.3.4, 348-349 and 2.3.2.6, 410-411. See Margaret Watkins Tate, “Obligation, Justice and the Will in Hume’s Moral Philosophy”, 100.
\(^{39}\) T 3.3.6.4, 619-620.
On Harris’ reading, Hume thinks the convention as a system is “beneficial to society”, and it is this benefit, and not any particular motive, that we approve of: “once it has this tendency, we naturally approve of it.”

Harris concludes from this that Hume’s view is that we approve of acts that accord with the convention of justice because we sympathize with the convention’s benefit to its members, and we approve of such actions regardless of the motive to act in this way. As Harris states, “it is in terms of their consequences, not in terms of their motives, that just actions are morally appraised in the first instance.”

Although Hume thinks we approve of acting out of a sense of duty, “in the case of justice, there is no one kind of motive which actions are taken to be prompted by, and which explains why these actions are approved of.”

The approval of actions that accord with the convention is derived from the approval of the system’s consequences, rather than approval of the just person’s motive.

The method of supplying Hume with an approved motive, which Garrett and Sayre-McCord adopt, is the most straightforward way to fit Hume’s account of justice into his account of the virtues he gives, and to which he refers at the beginning of his chapter on justice.

In support of this view, it seems unlikely that Hume mentions approval of motives at the beginning of the chapter on justice if he does not think that justice fits into his theory of the virtues. But it is a pressing question, as Harris insists, why Hume doesn’t offer such a motive himself if this was his intention.

Moreover, as we saw in the quote above, Hume clearly states that it is the

---

40 Harris, “Hume on the Moral Obligation to Justice”, 40.
41 Harris, “Hume on the Moral Obligation to Justice”, 27.
42 Ibid, 40. There must be a disposition to act justly, Harris explains, but there is no one motive that explains our approval of this disposition. Ibid, 40-41.
43 I do not mean to take a stance here on whether Cohon is right that it is a character trait, rather than a motive, that renders a person virtuous on Hume’s theory.
44 Harris, “Hume on the Moral Obligation to Justice”, 40.
approval of the consequences of the system of justice which leads us to approve of actions that accord with the convention.\footnote{Hume reiterates this point in Treatise 3.3.1.9, 577: “Now justice is a moral virtue, merely because it has that tendency to the good of mankind; and, indeed, is nothing but an invention to that purpose.”}

This leads me to conclude that both interpretations get something right about Hume’s account of moral judgment in the case of justice, but both are incomplete. From what we’ve seen, both approval of motive and approval of a system of conduct are found in Hume’s account of justice in the Treatise. These two approvals coexist in Hume’s account, offering a more complex picture of moral judgment with regard to justice. Hume thinks justice is a virtue of a \textit{person}, and that such a virtue is constituted by approval of a particular motive. Like the natural virtues, the artificial virtues are constituted by approval of a motive some people have which leads us to judge them virtuous. But Hume also thinks that we approve of systems of action because they benefit society, and thus we approve of acting according to the convention irrespective of motives. This dual description of moral judgment that appears in Hume’s account of justice thus limits the role of approval of motives to a subset of moral judgments. On this account, approval of particular motives is constitutive of the virtues, but approval of motives is not constitutive of approval of systems of conduct.\footnote{Cohon understands Hume to be saying that we can approve of actions that accord with the convention regardless of their motive. See Cohon, \textit{Hume’s Morality}, 171, 175-176. For support of this view see also T 3.2.5.4, 517, where Hume discusses approval of actions and character traits, without mentioning approval of motives.}

My suggestion that Hume holds both that justice is a virtue of a person and that actions that are required by the convention are approved of independently of there motive is not necessarily objectionable to those who suggest that Hume is offering a moral virtue theory. According to some virtue theory interpretations, we can distinguish between Hume’s view of
virtue and Hume’s view of right action. While virtue, on this reading of Hume, is defined by approval of motives, Hume can hold that right action is not decided by approval of a motive, but rather by what a virtuous person would do.\textsuperscript{47} On this reading, Hume holds that we approve of actions that accord with the convention because they are the kind of actions a just person would perform, irrespective of the motive to act in this way.

But notice that such a virtue theory would be more radical than my own view, because it would claim that for Hume, there is a right action that is approved of regardless of its motive in the case of all virtues, artificial and natural alike. Thus a person who acted in a way that was kind would act in a way that was right, even if they were not motivated by beneficence and did not have this virtue. This would require textual evidence to show that Hume thought there are right actions in the case of the natural virtues. My view does not go so far. Based on the textual evidence I have brought, it is particular to the artificial virtues that we approve of actions that accord with the convention regardless of their motive. I return to this distinction between the natural and the artificial virtues on this point in the last part of this chapter.

2. The Unity of Hume’s Explanation of Moral Judgment

The two kinds of moral approval that I argued appear in Hume’s account of justice could be thought to be too divergent to offer a unified source of moral judgment. In this part I argue that these different kinds of moral approval can be part of a unified theory of moral judgment, if we view them in light of Hume’s explanatory project in the Treatise. I argue that

\textsuperscript{47} I thank an anonymous reviewer for this suggestion.
Hume’s sentimentalist account of moral judgment, on which moral judgment is explained by the causes of moral sentiments, can unite approval of motives and approval of systems. This is because the sentimental theory is explained by the causes of the moral sentiment, agreeableness and utility to self and others, which are characteristics of both motives and systems of conduct.

It is curious that Hume does not claim to present a virtue theory or a natural law theory. Hume’s professed aim in his discussion of morality is to explain the human psychology that underlies it. Hume’s discussion of morality is part of his general explanatory project, which he lays out in the introduction to the Treatise: “There is no question of importance, whose decision is not compriz’d in the science of man… In pretending therefore to explain the principles of human nature, we in effect propose a compleat system of the sciences, built on a foundation almost entirely new, and the only one upon which they can stand with any security.”

Hume’s aim is to place our knowledge on solid ground, by pointing to the psychological principles that give rise to it. In the case of morality, too, it is not Hume’s main goal to offer a moral theory. Rather his goal is to explain what the psychological principles are that give rise to our moral judgement.

I would like to emphasize that I do not mean that Hume’s explanatory project has no normative implications. Scholars disagree about whether Hume’s explanatory project is descriptive or normative. Michael Gill argues that Hume saw himself as describing the “Springs & Principles” that stand behind moral actions and moral judgments. Other important scholars think that Hume’s explanatory project also offers justification for engaging in our moral

48 T Introduction 0.6, xvi.
practices. In the following I take a third position. I agree with Gill that Hume does not aim to justify the ways in which we make moral judgments, but I think Hume is not being merely descriptive. When Hume offers an explanation of how and why we come to acquire moral concepts and make moral judgments, he describes a normative standard. Hume’s explanation of the source of moral concepts has normative import, since concepts are themselves normative. Just as an explanation of what a virtue is dictates what can and what cannot be called a virtue, an explanation of why and how we come to acquire certain moral concepts utilized in moral judgment tells us when we are succeeding and when we are failing to make moral judgments.

As mentioned above, Hume argues that we can explain our judgements of virtue if we look at the approval of the motives of individuals. But in order to understand the place of approval of motives in Hume’s theory of moral judgment, we must view this statement in context of his method of explaining moral judgment. Hume lays out his method for explaining moral judgment in the following passage:

Now since the distinguishing impressions, by which moral good or evil is known, are nothing but particular pains or pleasures; it follows, that in all enquiries concerning these moral distinctions, it will be sufficient to shew the principles, which make us feel a satisfaction or uneasiness from the survey of

---


51 Other philosophers have argued that Hume has both descriptive and normative goals in view. See Kate Abramson, “Happy to Unite, or Not?”, Philosophy Compass, 1 (3), 2006, 290-302; Jacqueline Taylor, Reflecting Subjects: Passion, Sympathy and Society in Hume’s Philosophy, Oxford University Press, 2015.

52 For a discussion of Hume’s ethical theory as deriving from his explanatory project see Daniel E. Flage, “Hume’s Ethics”, Philosophical Topics, 13 (3), 1985, 75-88.

53 See Cohon, Hume’s Morality, 161.

54 Flage is helpful on this: explanation is an induction, while a justification is a deduction. Hume’s project is primarily an inductive one, but once he acquires concepts of virtues or of moral judgment through induction these can justify particular instances of moral judgment through deduction. See Flage, “Hume’s Ethics”, 75-78.
any character, in order to satisfy us why the character is laudable or blameable. An action, or sentiment, or character is virtuous or vicious; why? Because its view causes a pleasure or uneasiness of a particular kind. In giving a reason, therefore, for the pleasure or uneasiness, we sufficiently explain the vice or virtue.55

Hume starts out by saying that he aims to explain “moral good or evil” that are known by certain impressions. But Hume is not content to describe our moral concepts as derived from certain sentiments. In his “enquiries concerning… moral distinctions” Hume wants to “shew the principles, which makes us feel a satisfaction or uneasiness”. In order to explain moral judgment, we point to the principle or cause that brings about the moral sentiments. Hume then uses vice and virtue as an example: he says it is only “in giving a reason, for the pleasure or uneasiness from the survey of any character” that we answer the question “why the character is laudable or blameable”. Hume thus thinks that moral judgment of vice and virtue can only be explained by pointing to the causes of the moral approval and disapproval we feel in surveying a person’s character. By reducing from different moral judgments to a common moral approval or disapproval, Hume is able to search for the common principle, or cause, of the moral judgments. These causes, Hume concludes later, are agreeableness or usefulness to self or others.56 Finding the common cause then explains why we make the moral judgments that we do.57

55 T 3.1.2.3, 471.
As this quote shows, Hume’s aim is to find an explanation for moral judgment, and for Hume the explanation will be in terms of causes. But this aim that Hume sets as the goal of his enquiry in morals should lead us to view Hume’s search for the approved motives of the virtues as a first step in the search for the causes of our approval. Observation of the judgement of virtues allows Hume to conclude that it is the approval of motives that constitute the virtues, but this for him is only a first step, on the way to the induction that shows that the motives approved of share certain characteristic that explain the judgment. Thus in utilizing this methodology Hume is interested in our judgement of virtue and vice as instances of moral judgement. His conclusion regarding the causes of moral judgment in the case of character does not exclude the possibility that there are other kinds of moral judgment in addition to judgments of character.

My suggestion is that when it comes to justice, Hume does not change his methodology. Rather he observes that we have another kind of moral judgment that is different from judgement of character. Hume leaves open the possibility that looking for the approved motive of justice can explain the virtue of the just person. But Hume thinks approval of a motive on its own does not cover all our approvals in the case of justice. This is because Hume observes that we approve of actions that are in compliance with the conventions of justice, even when these actions are not motivated by an approved motive and do not come together with approval.

---

58 I get this idea that Hume is interested in the virtues as experiments from Norman Kemp-Smith. He writes that “the Treatise in all its parts is devoted to testing and confirming, by selected ‘experiments’ of the hypothesis, suggested by his [Hume’s] own studies in ethics, that not reason but nature, not knowledge but feeling and instinct, are the ultimate controlling forces in all the various domains… of human experience,” Norman Kemp-Smith, The Philosophy of David Hume, Palgrave MacMillan, New York, 1941, 62.
of a person. He observes that we approve of actions that are in accordance with the convention of justice independently of their motive, and when he looks for a cause for this approval he finds that we approve of them because they contribute to a system of actions of which we approve because of its utility to society. As Hume says in his summary of his account of justice, “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.”

This interpretation of Hume is compelling if we think of the phenomena Hume might be observing. My claim is that empirically, the phenomena of justice does in fact include both judging a person just, because of a particular motive, and approving of actions regardless of their motives. We admire and judge some people to possess the virtue of justice. As we saw, what the particular motive that renders a person just is controversial. But putting aside the question of which motive is approved of in the case of justice, we can acknowledge that there are instances in which we are aware that a person returns a loan while she was in difficult financial straits, out of either respect for the law or doing her share, which brings about approval of her motive and cause us to judge her a just person. At the same time, we approve of returning a loan, irrespective of motive, since we think it contributes to the functioning of a healthy society. Approval of this action does not depend on its being motivated by any particular motive. We approve of this action even when it is motivated by fear or habit, which are not motives we particularly approve of and do not render a person virtuous. Although we approve of this action, we do not particularly admire the person who performs it. Thus while there is a virtue of justice,

---

59 T 3.2.6.11, 533-534.
a character trait we approve of in people who are motivated by regard for the law, we also approve of just actions without this motive.\textsuperscript{60}

It is the observation of these two kinds of moral approval in justice that led Hume to develop his distinction between the natural and the artificial virtue, which is meant to give a full account of the phenomena of moral approval we observe when it comes to justice. Hume thinks that we approve of the just person because of our approval of some motive. He also thinks that we approve of actions that are in accordance with the convention, because we approve of the outcome of a just system of conduct. Both of these kinds of moral judgment found in the case of justice are part of Hume’s theory of moral judgment. What unifies them is Hume’s sentimentalist theory of moral judgement, which he claims has its source in approval that is explained by the agreeableness or utility to self or others.

On this interpretation, approval of motives is neither the main focus of Hume’s account of moral judgment nor is it irrelevant to it. Since Hume’s goal is to find the general causes of moral judgment, he is interested in the various kinds of moral judgment. Hume observes that the majority of judgments we make are of individuals’ characters, and thus approval of motives takes up a large part of his enquiry. As he says, virtues “form the most considerable part of morality.”\textsuperscript{61} But Hume’s moral theory is not limited to his account of the virtues, since he finds that we don’t only judge people’s characters. In the case of actions that

\textsuperscript{60} Compare Hume’s view that it is the utility of justice that renders it virtuous in the Enquiry, M 3.13, 188: “Hence justice derives its usefulness to the public: and hence alone arises its merit and moral obligation.” See also “The necessity of justice to the support of society is the SOLE foundation of that virtue; and since no moral excellence is more highly esteemed, we may conclude, that this circumstance of usefulness has, in general, the strongest energy, and most entire command over our sentiments... it is the SOLE source of the moral approbation paid to fidelity, justice...” M 3.48, 203-4.

\textsuperscript{61} T 3.3.6.2, 618-619.
comply with the convention of justice, we also approve of them because of their contribution to a system we approve of.

The unity of Hume’s theory does not depend on fitting justice into a virtue or a natural law theory. The unity of his theory depends on being able to explain various kinds of moral judgment with the causes he has pointed to. Hume is interested in the approval of systems of conduct (and places it first in his account of the virtues) because it might challenge his explanation that moral judgment has its source in approval caused by agreeableness or usefulness to self or others. But his explanation stands because Hume shows that actions that accord with a convention of justice are approved of because of their contribution to a system that is useful to self and others. In fact Hume’s theory of moral judgment is shown to be stronger, because it points to causes that can explain different kinds of moral judgment.62

One passage that challenges my reading is the first passages of the chapter on justice, where Hume seems to insist that we do not judge actions virtuous in any way other than by looking to their motive. As he says:

’Tis evident, that when we praise any actions, we regard only the motive that produced them, and consider the actions as signs or indications of certain principles in the mind and temper. The external performance has no merit. We must look within to find the moral quality. This we cannot do directly; and therefore fix our attention on actions, as on external signs. But these actions are still considered as signs; and the ultimate object of our praise and approbation is the motive, which produc’d them.63

---

62 For the discussion of the strength of a theory depending on its being able to explain dissimilar phenomena, see Yoram Hazony, “Induction in Newton’s Methodology”, manuscript, 21-22. Hazony shows that the generality of a cause, according to Newton, is not the number of instances it encompasses but its ability to explain phenomena that are dissimilar.

63 T 3.2.1.2, 477.
If we read Hume’s claim here as a claim regarding all moral judgment of actions, it contradicts my suggestion that Hume thinks there is approval of actions that accord with the convention of justice which is derived from approval of the system of actions. I therefore suggest to read this claim as one pertaining to virtue and vice, and not to all moral judgment. Since I read Hume as maintaining that there is a virtue of justice of a person, this passage should be read as an introduction to the enquiry into the *virtue* of justice, and not as a general claim regarding all moral judgment. Hume begins his account of justice with an explanation of the approval of a just character, but he also goes on to explain the approval of the just system. In support of my reading, notice that Hume does not argue in this passage that we judge actions according to their motives and not in light of a system they contribute to. He only dismisses judging the actions on their own, since “the external performance has no merit”.

A similar passage that poses a challenge to my reading is found later in the Treatise, when Hume turns to discuss the natural virtues. Here again Hume seems to emphasize that in our study of morality we should not look for approval of actions at all.⁶⁴

> We are never to consider any single action in our inquiries concerning the origin of morals; but only the quality of character from which the action proceeded. These alone are *durable* enough to affect our sentiments concerning the person. Actions are, indeed, better indications of a character than words, or even wishes and sentiments; but ‘tis only so far as they are such indications, that they are attended with love and hatred, praise or blame.⁶⁵

Hume seems to be saying that actions are never the direct object of our approval, but are only judged good or bad derivatively, by the approval of their motive. This seems to contradict my

---


⁶⁵ T 3.3.1.5, 575.
claim that in the case of justice, Hume thinks that there is both approval of a person, explained by the approval of the person’s motive, and approval of actions, explained by approval of the system they contribute to.

There are two ways to respond to this challenge. First, Hume may not be completely consistent. It may be that in his account of justice he shows that we approve of a system of actions because of its benefit to society, while here he claims that we only approve of actions because of their motive. But there is also the possibility that the two claims are reconcilable. Notice that Hume says here that in our enquiries concerning morality we should not consider “any single action”. He explains why this is – he thinks that single actions do not elicit moral approval or disapproval, since they are not “durable” enough to excite that kind of response. It is unclear whether Hume means that single actions are too transient to influence us, or whether he is saying that single actions can have different consequences in different contexts and therefore they don’t elicit consistent moral judgement. Whichever one it is, Hume is claiming that we don’t approve or disapprove of individual actions. This explanation leaves open the possibility that actions that are part of a “combination of men, in a system of conduct” do excite approval, because they are part of a system which gives them a “tendency” to benefit others and the durability needed to excite approval. Although Hume says here that qualities of character and not actions “affect our sentiments concerning the person”, this is compatible with the view that we also make other moral judgments of systems of action.

In this part I defended the view that Hume theory of moral judgment includes both approval of motives that give rise to virtues and approval of systems of conduct that give rise to

---

66 T 3.3.6.4, 619-620.
approval of actions irrespective of motive. I argued that what unifies Hume’s pluralist account of moral judgment is the causal explanation Hume gives for it: both kinds of approval are explained by agreeability or usefulness to self or others. I suggested that while Hume thinks we need durability to elicit moral approval, and this is why a single action does not elicit moral approval, such durability is not unique to character and also characterizes systems of conduct.

3. Contribution of Non-Approved Motives to the Constitution of the Artificial Virtues

So far I’ve followed the scholarship on Hume in discussing approval of motives, which constitute virtues, and approval of systems of conduct, which does not arise from approval of a motive. In this part I turn to discussing the role of non-approved motives in Hume’s account of moral judgment. I show that Hume observes that non-approved motives, and particularly self-interest, have a contributing role in judging that justice is a virtue. Non-approved motives help maintain the convention, which is a pre-requisite for the approval of the motive that constitutes the virtue of justice. I claim that this contributing role of non-approved motives is distinctive of the moral judgment of artificial virtues.

As we saw exemplified in the discussion between Harris, Garrett and Sayre-McCord, the interest in motives in Hume scholarship is mostly an interest in approved motives. Because Hume argues that virtues are constituted by approval of motives, it seems that only approved motives are of any relevance to Hume’s theory of moral judgment. The exception to this is that it is now commonplace to distinguish between Hume’s genetic question and his moral question, when it comes to justice, as we saw above. Thus it is agreed that Hume thinks that self-
interest, although not a particularly approved motive, is the motive for instituting the convention of justice. This non-approved motive has an important role in bringing about the convention of justice, but it is not thought that it is relevant to the moral judgment that we make, when we approve of the motive of the just person. This is because self-interest is viewed as central only to the institution of the convention, while the approval of the motive of the just person occurs only once this convention is in place.

But surprisingly, Hume displays a continued interest in the non-approved motives people have to act according to the convention of justice even once it is up and running, and especially to self-interest as such a motive. It seems that for Hume, non-approved motives to act in accord with the convention of justice are of relevance to his account of moral judgment in the case of justice. Let’s first look at the textual evidence for this, and then consider what role non-approved motives have in Hume’s account of moral judgment in the case of the artificial virtues.

In 3.2.1 Hume considers and discards a number of natural motives which we could think are the approved motive of the virtue of justice. Hume first considers and discards self-interest. Hume ends this discussion with the following: “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.” With the negation “no real or universal” Hume concludes that there is no natural, and no single, motivation for “observing the laws of equity”. Thus in addition to

---

67 For example, Knud Haakonssen points out that Hume’s discussion of self interest as the motive to join and found the convention emphasizes that the benefits of justice go well beyond the intentions of those who instituted the convention. In order to show this, Hume must show that the convention was instated because of self interest, but benefits society as a whole. Knud Haakonssen, The Science of a Legislator: The Natural Jurisprudence of David Hume and Adam Smith, Cambridge University Press, 1981, 20-21.
68 T 3.2.1.10, 480.
69 T 3.2.1.17, 483.
the conclusion that there is no approved natural motive of the virtue of justice, Hume also concludes that there is no single motive to act in accordance with the conventions of justice. This conclusion seems superfluous. Why would it matter that there is no universal motive to comply with the convention of justice? Clearly there are many motives to act in any virtuous way, but these are of no interest to someone who is interested in moral judgment of virtues, which as we saw is the enquiry into approved motives. (Notice that there is no similar conclusion in the case of the natural virtues. We may have all kinds of motives to act in the way that a beneficent person would act, but since these do not render us virtuous they are irrelevant to Hume’s discussion of beneficence.)

We would think that once Hume shows that self-interest cannot be the natural, approved motive of the virtue of justice, he would have no more reason to discuss it. But Hume returns to discussing self-interest as a motive to act in accordance with the convention of justice once it is up and running in 3.2.2. Hume first mentions that in small societies, self-interest is sufficient to motivate people to follow the convention: “To the imposition then, and observance of these rules, both in general, and in every particular instance, they are first mov’d only by a regard to interest; and this motive, on the first formation of society, is sufficiently strong and forcible.”70 Hume then says that in large societies self-interest doesn’t usually motivates us to act in accord with the convention of justice, where the interest we have in following the convention is far removed: “when society has become numerous, and has encreas’d to a tribe or nation, this

---

70 T 3.2.2.24, 498-500, emphasis added.
interest is more remote; nor do men so readily perceive, that disorder and confusion follow upon every breach of these rules, as in a more narrow and contracted society.”

But although our blindness to our far-removed interest makes it difficult to be motivated by self-interest to act according to the convention of justice, other forms of self-interest are developed, that make it in our immediate interest to comply with the convention. Maintaining our reputation is in our self-interest, and Hume claims that maintaining our reputation mostly coincides with acting according to the convention of justice and therefore consistently motivates us to act in this way. “There is nothing which touches us more nearly than our reputation, and nothing on which our reputation more depends than our conduct, with relation to the property of others.” In addition to this commonplace self-interest ed motivation to act in accordance with the convention, Hume also thinks there are some wealthy individuals whose direct self-interest it is to maintain the conventions of private property. These are the people, Hume suggests, who seek to put in place solutions to make compliance more immediately in everyone’s self-interest, by instating punishment for transgression. Once this punishment is instated, self-interest becomes a common motivation for acting in accordance with the conventions of justice.

This string of comments is strange, in light of Hume’s conclusion that interest is not the approved motive of justice and Hume’s view that virtues are constituted by approved

---

71 Hume goes on to say that although our self interest in the existence of the convention of justice does not motivate us to act in accordance with it, we do become aware of this interest when others don’t follow the convention in their transactions with us: “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,” T 3.2.2.24, 498-500.
72 T 3.2.2.27, 501.
73 T 3.2.7.6-7, 537-538.
motives. Why does Hume mention that there is no single motive to act according to the convention, and why does Hume discuss how self-interest motivates us to act according to the convention once it is up and running, if virtues are sufficiently explained by an approved motive and self-interest is not this approved motive?

My account of Hume’s theory of moral judgment sheds light on why Hume continues this discussion of the role of self-interest. As I argued in part 2, we must read Hume’s moral theory in light of his explanatory project. I proposed that Hume views his discussion of virtue as a case in which he utilizes his explanatory method to progress in our knowledge of morality. I argued that when it comes to justice, Hume thinks that approval of a motive is not sufficient to explain our moral judgment, and he therefore adds approval of a system of conduct as another kind of moral judgment we find with regard to justice. But now we can say something more. Even the virtue of justice is not fully explained by approval of a motive, on its own, as is the case for the natural virtues. The virtue of justice can only be fully explained if we take the existence of a convention into account. On their own, the actions of the just person could be “pernicious in every respect.”\textsuperscript{74} The movie of the just person is only approved of, and the just person is only judged virtuous, if there is a convention in place. Thus even for those who understand there to be a virtue of a just person who is motivated by a particular motive, this motive (be it rule following or doing one’s fair share) is only approved of in the context of a convention.

This does not yet answer our question regarding the role of non-approved motives in the judgment of the artificial virtues. In order to understand Hume’s interest in self-interest as

\textsuperscript{74} T 3.2.2.22, SBN 498.
a motive to act according to the convention, we must think more about the role of motives in constituting the convention. A convention only exists if it is upheld by most members of a society. This is because people are only motivated to limit the avidity for possessions if they rely on others to do so as well. As Hume emphasizes,

\[ \text{… this may properly enough be call’d a convention or agreement betwixt us… since the actions of each of us have a reference to those of the other, and are perform’d upon the supposition, that something is to be perform’d on the other part.}^{75} \]

Convention thus requires the supposition of general compliance. But supposing that others will comply with the convention does not at all require that others be motivated by the approved motive and be virtuous people. If a convention depends on the supposition that others are going to act in accordance with it, then all the possible motives, including non-approved motives, to act according to the convention contribute to the constitution of the convention. If non-approved motives contribute to the constitution of the convention, this means that non-approved motives contribute to the judgment that someone has the virtue of justice, which requires the existence of a convention. Because general compliance is necessary to judge the just person virtuous, the motives of all those who act according to the convention, and not only the movie of the virtuous person, are partly constitutive of the convention that is necessary for our judgment that the just person is virtuous.

---

75 T 3.2.2.10, 490. See also “tis only upon the supposition, that others are to imitate my example, that I can be induc'd to embrace that virtue; since nothing but this combination can render justice advantageous, or afford me any motives to conform my self to its rules.” T 3.2.2.22, SBN 498.
We must be careful to stress that self-interest is not the approved motive of the virtue of justice. Hume says that self-interest is neither an approved nor a disapproved motive, and thus it cannot give rise to virtue or vice. If the virtue theory is partially correct, then there is some other motive that we approve of that renders justice a virtue. But without deciding what this motive is, we can see that Hume thinks that the existence of other, non-approved, motives contribute to the existence of the convention, which is a pre-requisite for the approval of the motive that constitutes the virtue of justice. This is because non-approved motives have a constitutive role in maintaining the convention that is approved of in the case of justice. We may be able to imagine a society whose members were all just, and therefore always acted according to the convention of justice out of the approved motive. But Hume’s insight is that we in fact do not live in such societies, and that in the societies we live in, the virtue of the just person is not only dependent on the self-interested actions of those who instituted the convention of justice, but also on the self-interested actions of those who comply with the convention and thus maintain it.

This points to a substantial difference between the moral judgment in the case of the natural virtues and the moral judgment in the case of the artificial virtues. First, let us look at how the distinction between natural and artificial motives is usually understood. Garrett and Sayre-McCord maintain that artificial virtues are constituted by the disposition to act on a single, artificial motivation of which we approve, which renders the trait in question a virtue. On their reading, the difference between the natural and the artificial virtues is in the kind of motivation we approve of. In the case of the natural virtues we approve of a natural motive, whereas in the

---

76 T 3.2.2.14, 492-493.
case of the artificial virtues we approve of an artificial motive that develops in response to a
convention. 77

On Harris’ interpretation, justice is an artificial virtue because of the way it
achieves its useful consequences. As we saw earlier, Hume thinks that the approval of the
benefits of justice is natural. Indeed he emphasizes that if we did not have this natural approval
of its benefits, no convention could have taught us to approve of particular just actions. But these
benefits are brought about by the “combination of men, in a system of conduct”: they are not the
result of a single action, but rather of the action in conjunction with actions of other people. 78
The approval of justice is natural, but the way the approved consequences are brought about is
artificial because it depends on a system of actions and not merely on the motive of the action
being judged.

I think there is room to interpret Hume as pointing out both these distinctions, but I
wish to point out an additional aspect of Hume’s artificial virtues: the moral judgment of the
natural virtues is constituted only by the approved motive, while the moral judgment of the
artificial virtues is dependent on the existence of a convention, which is partially constituted by
non-approved motives. On my reading, the natural virtues are always motivated by the motive
we approve of. Thus actions that indicate that a person is kind will be motivated by beneficence.
There is a coextension between the actions motivated by a certain motive and the actions we
approve of. Since actions are only approved of derivatively, because of their motive, actions not
motivated by the approved motive only look like virtuous actions. 79 A person who gives charity

---

78 T 3.3.1.12, 579-580.
because she wants to be loved by her friends only \textit{seems} beneficent, but in fact is not. Thus in the case of the natural virtues, non-approved motives have no role in explaining the judgment that someone is virtuous.

In the case of the artificial virtues, on the other hand, there are many acts that are in accordance with the convention and thus contribute to the existence of the convention, that are not motivated by the approved motive. These actions are not virtuous actions, since they are not motivated by an approved motive. But they do contribute to the convention, which is a pre-requisite of the judgment that the just person is virtuous. Not only are these actions approved of, because of their contribution to a useful system of conduct, they also contribute to the constitution of the convention which is a pre-requisite for us to judge the just person virtuous.

This reading is supported by Hume’s explanation of the distinction between the natural and the artificial virtues. Hume explains the distinction as follows:

\begin{quote}
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, and is the object of some natural passion: Whereas a single act of justice, consider’d in itself, may often by contrary to the public good; and \textquoteleft tis only the concurrence of mankind, in a general scheme or system of action, which is advantageous.\footnote{80 T 3.3.1.12, 579-580.}
\end{quote}

While Hume implies that he will offer one distinction, “the only difference”, between the natural and the artificial virtues, he seems to point out two differences here: first, the way in which the good is achieved (from every action or from a system of actions), and second, the kind of motive that is approved (natural or artificial passion). We can see where the virtue theory and the natural law theory each get their interpretations of the difference between the natural and the artificial
virtues. But Hume seems to give more room and weight to the first explanation, and to view the question of the kind of motive as subsidiary. Hume says that in the case of the natural virtues, there is a good that comes from every single action, whereas in the case of the artificial virtues the good is achieved through the joint efforts of many people acting in many instances.

My suggestion is that the distinction in the way the good is achieved on the natural and artificial virtues also brings about a distinction in the role of non-approved motives in bringing about the moral judgment of virtue. In the case of the natural virtues good is brought about in every instance, and therefore that good can motivate every person because it is the “object of some natural passion”. We therefore morally approve of those people, and only those people, who are motivated by this good. Non-approved motives have no role in moral judgment in the case of the natural virtues. But in the case of the artificial virtues, good is not always immediately brought about by acting according to the convention. One’s actions contribute to a system that is overall good, but don’t necessarily bring about good on their own (and may do the opposite). This is why we approve of all actions that comply with the convention even when they are motivated by non-approved motives. But the outcome of this contribution of non-approved motives to the constitution of the convention is that they also contribute to the judgment that a particular person, with an approved motive, has one of the artificial virtues.  

If we could ensure that all people would be strongly motivated by an approved motive, then perhaps our judgment of the virtue of justice would not depend on other, non-approved motives. But since we cannot ensure that all people will have the virtue of justice, the

---

81 For discussion of the centrality of unintended consequences to Hume’s account of the virtue of justice, as well as the other artificial virtues, see Haakonssen, *The Science of a Legislator*, 19-20; Hardin, *David Hume: Moral and Political Theorist*, 48-51.
non-approved motives for complying with justice is central to Hume’s account of justice. The continued existence of the convention is partly constituted by these non-approved motives, and therefore the judgment that justice is a virtue is dependent on them as well.

**Conclusion**

Moral theorists often attempt to reduce our complex moral lives to one moral principle. Hume resists this temptation by pointing to four causes of moral approval: usefulness and agreeableness to self and others. But when scholars come to explain Hume’s moral theory, they often reduce his account of moral judgment to an approval of one kind. The virtue interpretation presents Hume as thinking that all moral judgment is a judgment of motives, while the natural law interpretation presents Hume as thinking that all moral judgment is a judgment of systems of conduct. I have shown that Hume maintains his pluralism not only with regard to the causes, but also with regard to the objects, of our moral judgment. In his account of justice Hume describes moral approval of both motives and systems, and thus Hume holds that there are different objects of moral judgment. What unifies this moral theory is not the objects of moral judgment, but the sentimental approval which is the source of these judgments and the causes that give rise to such a sentiment. On this view, approval of motives is only one form of moral judgment. This reading allows us to appreciate another aspect of Hume’s distinction between the natural and the artificial virtues. While the judgment of the natural virtues can be explained by the approval of a single motive, the judgment of the artificial motives cannot be explained in this way. The approval of a certain motive that constitutes the virtue of justice only comes about in the context of a convention. This convention, in turn, is brought about not only by the approved
motive but also by other, non-approved motives. Thus non-approved motives have a constitutive role in maintaining the convention, which in turn is a pre-requisite for the moral judgment of the virtue of justice.
Chapter 2: Humean Loyalty

In this chapter I offer a novel theory of loyalty utilizing aspects of Hume’s psychology. Hume discusses loyalty to government, but he doesn’t discuss loyalty as a character trait found in many forms of human relationships. This is a glaring lack in his moral psychology, since loyalty is a prevalent moral phenomena. But I believe that Hume’s moral psychology is uniquely suited for offering a psychological explanation of loyalty, the kind of explanation that is lacking in the literature on loyalty. Once we understand loyalty on Humean terms, I will argue, we are also in a better position to understand why loyalty is a virtue.

My analysis of Humean loyalty will start, in part 1, with a discussion of the account of self that arises from Hume’s discussion of pride. I argue that on Hume’s view the self extends to include others through association of ideas. In part 2 I suggest that loyalty is acting to fulfill the needs and interests of those who are part of the extended self because they are part of one’s extended self. In part 3 I emphasize the distinctiveness of this motive by showing how Humean loyalty is different from motives to act on behalf of others acquired through Hume’s mechanism of sympathy. This leads me to critique Hume’s account of the love of relations, in part 4, arguing that loyalty rather than sympathy explains these strong ties. In part 5 I argue that my theory of Humean loyalty offers a new explanation of the virtue of loyalty: loyalty is beneficial to others.

---

82 This discussion of loyalty is meant as a general of the trait, including loyalty to individuals, groups, nations and government. Hume does not discuss loyalty at any length, except for his discussion of allegiance. Hume uses the term loyalty at T 3.2.8.7, 545-546 and T 3.2.10.4, 556. It seems that Hume uses loyalty interchangeably with allegiance in this chapter, as well as in M 4.18, 210, where he seems to use fidelity to refer to women’s loyalty to their husbands (referring to the artificial virtue of chastity) and loyalty to refer to the artificial virtue of allegiance. See also A Dialogue 33, 335, “The most inviolable attachment to the laws of our country is every where acknowledged a capital virtue: and where the people are not so happy, as to have any legislature but a single person, the strictest loyalty is, in that case, the truest patriotism.” In countries governed by an individual, allegiance is understood as loyalty to a person or lineage rather than to the law, see T 3.2.10.15, 562-563.
because it offers a dependable motive to act on their behalf, and it is useful to the self because it motivates the individual to allocate time and resources to act on behalf of others, in a way that also strengthens the individual’s extended self. In the last part of this chapter I respond to the objection that Hume’s critique of factions implies that he did not think loyalty was a virtue.

1. Hume’s Extended Self

Scholars have discussed the role the opinions and sentiments of other people has in constituting the self on Hume’s theory.\(^\text{83}\) In this part I argue that there is another way in which other people constitute parts of the self on Hume’s theory. I show that Hume’s account of pride and humility discloses a view of the self as able to extend to include other people. This extension of the self creates a category of people who are psychologically and metaphysically separate from me (they have a separate body, a separate mind, separate passions and will) and yet they are a part of the self because they are associated with the other ideas that make up the self. This view of the self will later help explain the psychology of loyalty.

Hume’s second book of the *Treatise* is divided into two categories of passions, the passions of pride (and humility) and the passions of love (and hate). What distinguishes these passions are their objects\(^\text{84}\): the object of pride is the self,\(^\text{85}\) while the object of love is another person. The very same quality that causes us to love another person will cause us to feel pride

---


\(^{84}\) Hume calls this the subject, not the object, of pride. Hume distinguishes between the cause and the object of pride, “that quality, which operates, and the subject, on which it is plac’d.” As he says, “Both these parts are essential, nor is the distinction vain and chimerical,” T 2.1.2.6, 279.

\(^{85}\) T 2.1.2.2, 277.
when it is our own.\textsuperscript{86} What causes this change in passion is not an alteration in the quality, but rather in the object it is associated with. Hume insists that pride and humility can only have as their object the self: “When self enters not into the consideration, there is no room either for pride or humility.”\textsuperscript{87} Hume emphasizes the different objects of love and pride when he says that there is no real sense of self-love: “when we talk of self-love, `tis not in a proper sense, nor has the sensation it produces anything in common with that tender emotion, which is excited by a friend or mistress.” \textsuperscript{88} Love and pride are different categories of passions, distinguished by their objects.

Unexpectedly, the dichotomy between self and others emphasized in the distinction between the categories of pride and love is found to be imprecise. Hume claims that the causes of pride are not limited to qualities of the agent, but include the qualities of any other person related to the agent. Hume writes as follows,

To begin with the causes of pride and humility; we may observe, that their most obvious and remarkable property is the vast variety of \textit{subjects}, on which they may be plac’d. Every valuable quality of the mind, whether of the imagination, judgment, memory or disposition; wit, good-sense, learning, courage, justice, integrity; all these are the causes of pride; and their opposites of humility. Nor are these passions confin’d to the mind, but extend their view to the body likewise. A man may be proud of his beauty, strength, agility, good mein, address in dancing, riding, fencing, and of his dexterity in any manual business or manufacture. But this is not all. The passions looking farther, comprehend whatever objects are in the least ally’d or related to us. Our country, family, children, relations, riches, houses, gardens, horses, dogs, cloaths; any of these may become a cause either of pride or of humility.\textsuperscript{89}

\textsuperscript{86} T 2.2.2.3, 279. “Beauty, consider’d merely as such, unless plac’d upon something related to us, never produces any pride or vanity.”
\textsuperscript{87} T 2.1.2.2, 277.
\textsuperscript{88} T 2.2.1.2, 329-330.
\textsuperscript{89} T 2.1.2.5, 278-9.
Hume observes that as an empirical fact, pride and humility are not only caused by one’s own qualities, but also by the qualities of family members, and we could add perhaps also friends. We are proud not only of our own kindness or physical strength, but also of the kindness or physical strength of our parents, children, or countrymen.

This observation that we feel pride at the qualities of other people related to us does not lead Hume to change his mind about pride being directed to the self. Hume emphasizes that “‘Tis always self, which is the object of pride and humility; and whenever the passions look beyond, ‘tis still with a view to ourselves, nor can any person or object otherwise have any influence upon us.” But how can this be, that we feel pride and humility only “with a view to ourselves”, while we seem to feel pride not only in ourselves but also in those related to us?

In order to explain why we feel pride in the qualities of those related to ourselves, we must turn to Hume’ view of the self. Hume’s view of the self is famously referred to as the “bundle theory”. Hume opposes the view that the self is a metaphysically unified substance, and argues instead for the view of the self as a bundle of impressions and ideas connected together through their association with, and causal influence on, one another. Even our body, which people so clearly think of as part of themselves, Hume thinks is part of the self through association of ideas. Hume reiterates this view of the self in the opening paragraphs of the chapter on pride, so it is clear that it is this idea of the self that is the object of this passion: “Tis

---

90 T 2.1.3.2, 280.
91 It is worth noting that this is not Hume’s preferred term. He claims that the “true idea of the human mind” is best compared to a republic, since it “is a system of different perceptions or different existences, which are link’d together by the relation of cause and effect” (T. 1.4.6.19, 261). For discussion see Alanen, “Personal Identity, Passions, and ‘The True Idea of the Human Mind’”, 16-19.
92 See T 1.4.6.3–4, 252; Abstract, 28.
93 T 1.4.6, 251-254.
evident, that pride and humility… have yet the same OBJECT. This object is self, or that
succession of related ideas and impressions, of which we have an intimate memory and
consciousness.”

We will want to understand here an aspect of this idea of the self as it arises from Hume’s discussion of pride.

Amelie Rorty argues that on Hume’s theory pride discloses to us the contours of the bundle that makes up the self. She explains that the passion of pride is distinct from the passions of love or joy precisely because of how closely we relate its causes to ourselves, meaning how much the cause is part of ourselves. This delineation of the contours of the self is not consciously or voluntarily drawn. Rather the passions of pride and humility indicate which impressions and ideas make up the self. As Rorty explains, “Whatever it is that, as a given original fact, causes pride, turns out as a matter of fact to be what a person takes as significant and particularly close to her.”

Thus a person could enjoy or love an experience of watching the play of light on a leaf, but a person who feels pride in her enjoyment is a person who has this ability as a part of her idea of herself. On Rorty’s reading, therefore, pride points to what ideas make up the self:

“Hume has produced an account of the sources of the idea of the self characterized by the causes

---

94 T 2.1.2.2, 277. To some it may seem strange that I am using the view of the self in Book I in context of the discussion of self in Book II. Pitson argues that the view of the self Hume proposes in Book I of the Treatise is different from the view presented in Book II of the Treatise. See Pitson, Hume’s Philosophy of the Self, p. 1. Others have also held that Hume has two theories of the self, see Jane McIntyre, “Personal Identity and the Passions”, Journal of the History of Philosophy 27 (4), 1989, 545-557; Donald Ainslie, “Sympathy and the Unity of Hume’s Idea of the Self”, in Persons and Passions: Essays in Honour of Annette Baier, J. Jenkins, J. Whiting, and C. Williams (eds.), University of Notre Dame Press, Indiana, 143-173. I am not convinced you can make such a clear distinction. For the view that Hume only has one view of the self see Asa Carlson, “There is Just One Idea of the Self in Hume’s Treatise”, Hume Studies 27 (4), 2009, 35. Lilli Alanen emphasizes that the idea of the self in Book II corroborates the idea of the self in Book I. See Alanen, “Personal Identity, Passions, and ‘The True Idea of the Human Mind’”, 3-28. Alanen points out that the passions are already there in the account of Book I. Alanen, “Personal Identity, Passions, and ‘The True Idea of the Human Mind’”, 9.

of its various prides... At its initial level, the idea of the self is an idea of a class of ideas, composed of as many ideas as there are sources of pride.” 96

But pride does not only have this passive role. Both Rorty and Lorraine Besser argue that on Hume’s view the passion of pride does not only indicate to us what ideas make up the self, but also reconstitutes the idea of the self by prioritizing among these ideas. “Pride takes the items of biography and weights them.” 97 Through pride in qualities of objects related to ourselves, the self is reconstituted as having these objects as part of the self. As Rorty says, “It is a brute fact that pride, itself a pleasurable experience produces the idea of the self, the self as owner of a fine ancestral house, as mother of a fine young man.” 98 Because I find pleasure in the house of my ancestors, pride leads me to view myself as the owner of the ancestral home, for example. It is this reconstitution of the self through prioritizing among the impressions and ideas that come into our lives that turns the bundle of ideas into an agent. Only with the prioritizing of impressions can we be motivated to action. 99

Rorty’s discussion of the reconstitution of the self as an agent through pride is purposefully general, since people have all kinds of ideas as part of themselves. What I would like to develop here is something that she does not emphasize. Rorty mentions in passing that “When the idea of the self is enlarged or diminished, its passions and concerns for itself also

96 Ibid, 259-260. Rorty lists three other characteristics of the causes of pride that Hume discusses: they must be peculiar to ourselves, “relatively enduring” and valued by others. Ibid, 261-262.
97 Ibid, 263. Lorraine Besser argues that on Hume’s model we must distinguish between the self that causes pride and the self that is the object of pride. See Lorraine Besser, “Hume’s Practical Conception of the Self”, unpublished manuscript. Jacqueline A. Taylor mentions something similar in passing, see Reflecting Subjects: Passion, Sympathy and Society in Hume’s Philosophy, Oxford, Oxford University Press, 50.
99 Ibid, 264; Besser, “Hume’s Practical Conception of the Self,” 7-9. Since for Hume passions motivate action, scholars have suggested that we must look to this account of the self as it arises in book II to understand how the Humean self is motivated to action. See McIntyre, “Personal Identity and the Passions”; Rorty, “‘Pride Produces the Idea of Self’: Hume on Moral Agency”, 255-269.
change.” What is this enlargement and diminishing of the self that she mentions? If we think of ourselves as smart or stupid, good looking or ugly, we cannot say that there has been any enlarging or diminishing of the self. What does allow for the enlargement of the self is coming to view objects, and especially other people, as part of the self.

I suggest that it is this view of the self as extending to include others that can help us explain why it is that we feel pride and humility at others, while these are passions directed towards the self. As we saw, Hume observes that we feel pride and humility, rather than love and hate, in the qualities of those related to us. Hume says that pride can be caused by the qualities of “whatever objects are in the least ally’d or related to us”; “parts of ourselves, or something nearly related to us”; “in us or belonging to us.” But Hume doesn’t explain what it means that something is “part of ourselves”, nor what is the psychology of something belonging to us. As we saw, Rorty argues that pride and humility make clear to us what are the impressions and ideas that make up the self. But if qualities of other people who are related to us give rise to pride and humility, the ideas of these people must be part of the idea of the self that is the object of pride. This makes sense on Hume’s view of the self as a bundle, since the idea of other people can be associated with the self and thus become part of the idea of the self. Just as we associate the idea of our body with our self, and it becomes our body, we associate ideas of family and friends, team or country, with our self and they thus become ours. The effect of the association

101 T 2.1.2.5, 279.
102 T 2.1.5.2, 285.
103 T 2.1.8.6, 301.
104 Hume does not think that pride or humility depend on one achieving good qualities. Just as he thinks we approve of virtues and abilities, without a distinction made regarding our active role in bringing them about, so he thinks we can be proud of things we make our own or things that are our own by inheritance. For a critique of this position, see
is that we are proud of the strengths and good qualities of those associated with the self, while we feel humility with regard to their weaknesses, in the same way that we are proud of our own strengths and humbled by our own weaknesses.

Rorty thinks that the sense of mine that is relevant in the case of extending the self to include objects or groups is legally and socially defined, on Hume’s account.\textsuperscript{105} It is clearly true that for Hume, ownership of property is a concept that depends on socially developed conventions, and Hume emphasizes that we would have no such concepts without the convention. But I think that Rorty is too quick to group together ownership of property with the relation of others to the self, and therefore overlooks the fact that our self can extend to include others (individuals, objects or groups) independently of convention. The self that underlies the passion of pride can include other people or groups, I suggest, without any conventions in place that define these people and groups as legally belonging to the self.

Many (including Hume himself\textsuperscript{106}) have questioned the bundle theory of the self, since it is not clear what holds the impressions and ideas together and distinguishes them from other impressions and ideas. But what is important for our purposes here is that on Hume’s view the limits of the self are not set, since new impressions and ideas can be associated with the existing ones and become part of the idea of the self. This extension is not limited to impressions and ideas of our experiences or qualities, but also includes objects such as people and groups.

\begin{flushright}
\end{flushright}

\begin{flushright}
\textsuperscript{105} Rorty, ““Pride Produces the Idea of Self”: Hume on Moral Agency,” 259. Chazan also gives an important role to conventions and socializing, but for her their role is regulatory. They define when we are right to consider something part of ourselves, rather than merely fantasizing. See Chazan, \textit{Moral Self}, 24.
\end{flushright}

\begin{flushright}
\textsuperscript{106} T Appendix 20-21, 635-636.
\end{flushright}
One does not need to share Hume’s metaphysics of the self in order to observe that human beings have this tendency to enlarge their self to include other people. The self of a person who has a child expands to include this new addition to themselves – they are now a parent who has a particular child as part of their idea of themselves. The same can be seen when a person breaks off a long-standing romantic relationship – a part of themselves is severed or removed. Thus regardless of whether Hume is right that it is association that holds the bundle of ideas that is the self together, we can take from Hume’s theory the idea that the self can extend and contract to include others.107

There is some similarity between my suggestion that the self extends to include others and the view of the self that is developed by psychologists studying “oneness”.108 Some psychologists researching motivations to assist others have suggested that, at least some of the time, when we act on behalf of others we are not motivated altruistically but rather we act on behalf of others who we view as one with us and that we would think of in terms of “we”.109 This

---

107 The extension of the self to include others is discussed in the philosophy of love, see R.C. Solomon, Love: Emotion, Myth, and Metaphor, Anchor Press, New York, 1981; Roger Scruton, Sexual Desire: A Moral Philosophy of the Erotic, Free Press, New York, 1986; Robert Nozick, “Love’s Bond”, in The Examined Life: Philosophical Meditations, Simon & Schuster, New York, 1989, 68–86. These philosophers touch on a similar phenomenon to the one I am discussing, but they claim that it is particular to love. They therefore focus on the intentional extension of the self. On my Humean view, the extended self is an association of ideas which is not always intentionally or voluntarily brought about. Although this association may be caused by activities motivated by a certain passion, this is not the only way we come to have extended selves. I do agree with these authors that the extended self plays an important role in love, a component that I argue is missing from Hume’s account of love of relations, see part 4 below.


state, it is argued, can be distinguished from acting out of empathy with another, since the other is not viewed as other but rather as one with the self.

Although these studies are analyzing a similar phenomena to the one I describe as an extended self, they are missing something. The terms “we” and “onesss” seem to imply that there is a complete merging between the agent and the person who is part of the agent’s self. But when we feel pride in the wisdom of another person we do not stop viewing them as another person. Even when an agent acts on behalf of another, and even if the agent is thinking in terms of “we”, the distinction between the two people is not erased. This is why explaining these phenomena in terms of extending is more precise. This term retains the internal structure on which the narrow self, which has for example a certain body and a certain lie history, is more central to the idea of the self than other people who are associated with it. It conveys that the other person can be viewed as motivationally and metaphysically distinct, while constituting a part of the agent’s self.

2. Humean Loyalty

In this part I suggest that the Humean extended self described above offers the basis for a novel account of loyalty. The extension of the self to include others creates a narrow and an extended self, the latter including individuals who are distinct from the self metaphysically and psychologically but are part of the idea of the self. Humean loyalty is the character trait to tend to act on the motive to strengthen and benefit those who are part of one’s extended self. Loyalty is distinguished from self-interest but is a similarly strong motivation for action.
Loyalty is often characterized as the tendency to strengthen or support the interests of “one’s own”. As Andrew Oldenquist puts it, “When I have loyalty towards something I have somehow come to view it as mine. It is an object of noninstrumental value to me in virtue (but not only in virtue) of its being mine.”\(^{110}\) But what is this relation of mine, which generates a special motive to action? John Kleinig distinguishes “mine” as a non-causal relation, arguing that loyalty is directed towards objects constitutive of the self.\(^{111}\) But objects constituting the self seems nonsensical, on a view of the self as a unified substance or a soul with immovable limits.\(^{112}\) The view of the self as extending to include others can help us explain this. Because the limits of the self as a bundle of impressions and ideas is not set, on Hume’s account of the self, ideas of other people can constitute part of the self. Other people are constitutive of the self, and are therefore “one’s own”, when ideas of them are part of the idea of the self.

What is the motive to act on behalf of those who are part of the extended self, on Hume’s psychology? For Hume motives to action are passions, though not all passions motivate. Passions are reflective impressions, which means that they are impressions generated by the mind returning to ideas of pleasure or pain and responding to them.\(^{113}\) Thus passions are not derived from the impressions of sense, but rather are the mind’s original response to them. As Elizabeth Radcliffe points out, a Humean motive is not a complex of belief plus desire, what is


\(^{113}\) T 1.1.1.2; Radcliffe, *Hume, Passion and Action*, 17.
mistakenly called the “Humean model” of motivation. Rather passions themselves are motivating states: a passion is “a psychological state that could or might cause movement and action.” Motives are not reasons, because they are not justifications of actions. Rather they are the causes of action. But as Radcliffe notes, motives are potential causes, since they don’t always overcome other motives. Some passions are direct, like pain or hope, which means they respond to the impression of pain or pleasure alone. Other passions are indirect, which means they both have an impression that causes them and an impression or idea that they direct us towards, such as pride or love.

The most natural candidate for the motivating passion of loyalty is pride. I will consider this candidate, before I argue that loyalty is a distinct passion. As we saw, pride is felt at the good qualities of those who are part of the extended self, so it would make sense if pride also motivated us to act on their behalf. But does this fit with Hume’s account of pride? Lorraine Besser has argued that on Hume’s moral psychology a central motive to acquire virtue is pride-in-virtuous character traits, which consistently motivates us to acquire and strengthen virtues. Besser’s claim is that the pleasure of pride in one’s virtuous traits motivates us to acquire these traits. Interestingly, Besser only discusses the motivating influence of some of the Humean causes of pride, namely positive character traits of the self. This is because Besser is interested in the motivation to acquire virtue. I suggest that Besser’s claim that pride in virtuous

---

114 Radcliffe, Hume, Passion and Action, 16: “a motive for Hume is the passionate state, rather than a belief-desire complex”.
115 Ibid, chapter 1.
traits motivates action can be expanded to a claim regarding the motivating influence of other causes of pride: the qualities of people and groups who are part of the extended self. Since we feel pride in the good qualities of those people who are part of the extended self, and humility at their bad qualities, we are motivated to improve their qualities in the same way we are motivated to acquire virtues.

But this doesn’t seem quite right. We can think of examples where loyalty motivates a person to improve the qualities of the person they are loyal to: we can think of a teacher who is dedicated to the development of her pupil’s musical abilities. But many paradigmatic cases of loyalty are not cases of developing qualities of another person. Loyal people improve another person’s overall state or well-being, or take up an interest they have. Take for example a husband who gives up a good job and a comfortable home, along with proximity to family and friends, in order to move to another city so his wife can take up the job of her dreams as a professor. Although it would be natural for the husband to be proud of his wife professorship, it would be strange to claim that what motivates the husband to move is pleasure in his wife’s quality of professor. What seems to motivate the husband is a pleasure in his wife’s overall success and wellbeing, rather than in a particular quality that he is proud of.

In defense of this suggestion, it is possible that Hume is incorrect to characterize pride and humility as caused by qualities alone. Oldenquist points out that we feel pride not only in the qualities of another person, but also in their success and prosperity. “When I have loyalty toward something I have somehow come to view it as mine… I am disposed to feel pride when it
prospers, shame when it declines, and anger or indignation when it is harmed.” On Oldenquist’s view, pride and shame are not only caused by qualities, but also by the overall situation of those who are ours. Thus, for example, I feel the pleasure of pride watching my daughter start to crawl. The pleasure here is not in the quality of crawling, as is implied by Hume’s account, but in the advancement of the child. The same can be said for the loyal husband: he is not motivated by the pleasure of pride caused by the quality of professor his wife will acquire, rather he is motivated by the pleasure caused by her success and advancement.

But there is another problem with arguing that pride is the motive of loyalty, on Hume’s psychology. Besser’s view that Humean pride is motivating is controversial, since Hume sometimes seems to claim that pride is a passive passion which does not motivate action. Scholars argue that Hume’s account of pride is characterized by self-contentment, and benefits the proud person precisely by enabling him to find pleasure in what he already is and has. Such passivity could not motivate loyalty, so I suggest that we look elsewhere for the passion that motivates loyalty.

Hume claims that some of the passions that do not motivate are most often accompanied by a passion that does motivate. Thus Hume thinks the love and hate are passive

\[117\] Oldenquist, “Loyalties”, 175. For the relationship between pride and loyalty, see also Keller, Limits of Loyalty, 21; Kleinig, On Loyalty and Loyalties, 19.

\[118\] Taylor emphasize this tranquility found in pride as the basis for its virtue. See Jacqueline Taylor, “Hume on the Dignity of Pride”, Journal of Scottish Philosophy, 10 (1), 2012, 37-38. This is drawing on Hume’s T 2.2.6.3, 367, where he says that pride and humility are “pure emotions in the soul, unattended by any desire, and not immediately exciting us to any action”, “completed in themselves” and “rest in that emotion they produce”. Besser emphasizes that Hume does think pride can strengthen other sentiments that motivate, so that pride can be an indirect motive for seeking to acquire or strengthen causes of pride such as virtue. See Lorraine Besser-Jones, “Hume on Pride-in-Virtue: A Reliable Motive?”, 173-174. She relies here on Hume’s statement that pride gives “new force to our desire or volition, joy or hope,” T 2.3.9.4, 439. Annette Baier tries to ameliorate this tension by arguing that pride does not motivate us to acquire new possessions, but does motivate to conserve one’s possessions. Such a conserving motivation will, indirectly, motivate actions to inhibit harm. See Annette Baier, “Master Passions”, in Amelie O. Rorty (ed.), Explaining Emotions, University of California Press, Berkeley and Los Angeles, 1980, 404-407.
passions, but he claims they are “always followed by, or rather conjoin’d with” the motivating passions of beneficence and anger. A person who loves another will usually be motivated to benefit her, while a person who hates another will usually be motivated to harm her. I contend that the passion of loyalty similarly accompanies the passive passions of pride and humility.

One who is proud of or humiliated by those who make up his extended self will usually be motivated to improve their situation. Thus we find that just as the pleasure or possible pleasure in one’s own virtues can motivate action through the passion of pride, the pleasure or possible pleasure in the success and flourishing of those who are part of the extended self can motivate action through the passion of loyalty. The pleasure that causes the passion of pride is pleasure in a quality of the self, or of others who are associated with the self. The pleasure that causes the motive of loyalty is the success, strength and flourishing of those who are associated with the self. The pain that causes loyalty is the need, failure or weakness of these objects.

Although on this view loyalty is taking up the interests of someone who is part of one’s extended self, loyalty is distinguished from self-interest. Self-interest is the motive to strengthen the objects of pride and humility that are part of a narrow self. Loyalty is the motive to strengthen those who are metaphysically separate from me (they have a separate mind, separate passions and will) and yet are a part of me because they are associated with the idea of my self and are thus a part of my extended self.

---

119 T 2.2.6.3, 367.
120 This parallel to love and hate is not complete, since pride and shame are not accompanied by opposing motives of action. Ewin raises the possibility that we are loyal to those we are proud of. R.E. Ewin, “Loyalty and Virtues”, _Philosophical Quarterly_, Vol. 42, No. 69, Oct. 1992, p. 409. On my Humean account you can be loyal to those you are either proud or ashamed of, so long as they are associated with the self.
While the passion of loyalty is caused by the needs of someone else, loyal action also benefits the loyal agent. When the loyal person gives care, support or encouragement to another, the loyal person benefits because this strengthens his or her own extended self. The parts of the self cohere together, and as a whole this extended self is stronger and more able to face difficulties. When a person does not act out of loyalty when someone who is part of his extended self needs assistance, the cost is not only for the other person. The cost for the agent is that his own sense of self is shaken, and his ability to act in ways that will strengthened his own extended sense of self is questioned.

Extending the self through association of ideas does not erase the distinction between individuals. What the loyal person gets from taking responsibility for the interests of someone associated with himself is different from what he gives to the person he is loyal to. Let us take an example to clarify this important point. If a man quits his job to move to another city, because his wife has been hired for her dream job there, he gives and receives different things. His wife gets her dream job and a sense of support. He gains a sense of accomplishment, strength, and happiness. Since the consequences of the loyal action are different for the loyal person and for the object of loyalty, it is clear that loyalty is not caused by erasing the distinction between persons.

Because the extending self does not erase the distinction between self and others there will always be possible conflicts between loyalty and self-interest. The existence of a motive of loyalty does not prevent such conflicts, and indeed conflict is almost immediately

---

121 Baier thinks that Hume’s ethics rearranges our social situation so that there are no conflicting interests: “Where, on the more contractarian model, morality regulates and arbitrates where interests are opposed, on a Humean model… morality’s main task is to rearrange situations so that interests are no longer opposed,” Baier, Moral Prejudices, p. 70-71. I think loyalty leaves room for conflict.
associated in our minds with loyalty. A team member will be offered a better position with another team, or a soldier will endanger his life to protect his country. In these instances there are clear conflicts, where acting on the interest of someone associated with the self will harm the self in some way: staying with your team will cost you financially, defending your country may end your life.

What changes in recognizing the associations that make up the extended self, and being motivated by them, is one’s perspective on such conflicts. The loyal person views these as conflicts between the interests of a narrow and a broad self, and not as conflicts between self-interest and the interest of another person. To the loyal person, actions on behalf of the object of loyalty are not a sacrifice of her self-interest, because they forward the interest of her extended self. This does not eliminate conflict, but it changes it: it turns a zero-sum game into a reality where the agent can benefit while benefitting others. The extension of the self renders what would have been sacrifice of self-interest on behalf of one’s team or spouse a new kind of gain. Sticking with one’s team or moving with one’s spouse can benefit one’s extended self, even as it requires sacrifice of the narrow self.

This is what is difficult or scary about loyalty. People often think that loyalty is scary (or even wrong) because it requires giving up one’s self-interest. But as we saw, loyalty doesn’t require that a person completely overlook their own interest, but rather motivates a person to choose those interests that accord with the interests of others. Thus self-sacrifice is not what is

122 Keller also thinks that loyalty changes our beliefs, but he thinks that they change our judgment of the object of loyalty. What I suggest here is that the kind of belief that changes is very specific: it is beliefs concerning the existence of a conflict of interest. Beliefs concerning the qualities of the object of loyalty do not need to change because of the passion of loyalty, though the importance of the object to the self has changed.

123 For the view that Hume thinks this shared perspective is the moral perspective, see Gerald J. Postema, “Morality in the First Person Plural”, Law and Philosophy 14, 1995, p. 35-64.
difficult about loyalty. The hard part is the burden of being responsible for others, even when this doesn’t require completely overlooking one’s own interest. It is the very broadening of responsibility, having to think of what others need, and fearing that we cannot really help them that is the challenge in being loyal.

We can thus define loyalty as follows:

**Loyalty**: acting on behalf of the needs and interests of those who are merely part of the agent’s extended self when motivated by the fact that they are part of the agent’s extended self.

Loyalty is thus both a way of acting – it is acting on behalf of the interests and needs of others – and it is motivated by a particular motive – benefitting those who are part of the agent’s extended self. The causes of the motive of loyalty are the needs of someone who is part of the extended self. This means that if one acts on behalf of a friend or a child for selfish reasons, motivated by the benefits to oneself, one is acting as the loyal person would act but not being loyal.

There are two typical kinds of loyal actions. First, loyal people are motivated to do what is necessary to continue the relationship with their object of loyalty: couples don’t look for new partners, team members don’t look for a new team. A loyal person does not seek ways in which to sever the association between the object of loyalty and their self. This motivation of loyalty is not unlimited: we may find that we no longer have the passion of loyalty and are motivated to end a relationship. But such actions will not be motivated by the passion of loyalty, but by competing passions (self-interest, moral disapproval, fear or anger).\(^\text{124}\)

\(^{124}\) There are some rare cases where we might think that we are harming our object of loyalty, and in such instances, it may be that we decide that terminating the relationship is the loyal action to take.
Second, as we said the object of loyalty is one that can be prosperous or decline, and therefore can be strengthened or weakened. The loyal person is motivated to act in ways that strengthen the object of loyalty: parents feed and educate their children, lawyers defend their clients in court, friends help each other when they are sick or depressed. Loyalty is what turns these needs of others for food, education, defense or assistance, into pressures on us to act, as if they were our own needs. Although a parent is not hungry because his child is hungry, the hunger is pressing as though it is his own need. Although a friend’s depression is not the same as one’s own depression, this friend’s difficulty becomes one’s own difficulty to be addressed. Many loyalties have their actions dictated by the expectations of certain relationships. Thus they require that the loyal person take on specific interests of the person he is loyal to, and not all those interests. A loyal spouse is expected to not have a romantic relationship with any other person, while a loyal friend may have other friends.

3. Distinguishing Humean Loyalty from Humean Sympathy

Hume thinks we can acquire the sentiments of others through sympathy, and that sympathy varies with the closeness of a person to us, so we sympathize more strongly with those who are related to us. A Humean could think that strong sympathy is where we should look for Hume’s theory of loyalty.\textsuperscript{125} In this part I argue that loyalty is not strong sympathy. Sentiments

\textsuperscript{125} “The stronger the relation is betwixt ourselves and any object, the more easily does the imagination make the transition, and convey to the related idea the vivacity of conception, with which we always form the idea of our own person… The relations of blood, being a species of causation, may sometimes contribute to the same effect; as also acquaintance, which operates in the same manner with education and custom; as we shall see more fully afterwards. All these relations, when united together, convey the impression or consciousness of our own person to the idea of the sentiments or passions of others, and makes us conceive them in the strongest and most lively manner.” T 2.1.11.6-7, 318.
acquired through sympathy are a weak motivations to action on behalf of others, while loyalty is a strong motivation to act on behalf of some other people. This difference in motivation is explained by the different mental associations that give rise to each.

Hume argues that being influenced by the sentiments of other people is basic to our psychology. We are influenced by the pain of a passerby, or even the pleasure of a person we have not met, through our sympathy with them. “No quality of human nature is more remarkable, both in itself and in its consequences, that the propensity we have to sympathize with others, and to receive by communication their inclinations and sentiments, however different from, or even contrary to our own.”126 But Hume thinks sympathy is only weakly motivating to action: “My sympathy with another may give me the sentiment of pain and disapprobation, when any object is presented, that has a tendency to give him uneasiness; tho’ I may not be willing to sacrifice any thing of my own interest, or cross any of my passions, for his satisfaction”.127 The main consequence of sympathy is not motivation to act on behalf of others, but sharing in their sentiments. This sharing of sentiments is crucial to Hume’s sentimental account of moral judgment,128 but it cannot serve as his explanation of how it is that we are strongly motivated to act on behalf of our family, friends or countrymen.

126 T 2.11.2, 316. As Hume says in the Enquiry, “he must be more or less than man, who kindles not in the common blaze,” M 9.9, 275-6.
127 T 3.3.1.23, 586-587.
In order to understand why it is that sympathy is only weakly motivating to act on behalf of others, while loyalty is strongly motivating, we must look at what each relates to the self. Let us first remind ourselves of the explanation Hume gives for sympathy. Hume argues that in sympathizing we acquire the sentiment of another person by observing its outward effects. For example, we see a person yelling, turning red or acting violently, and the idea of anger rises in our mind. This idea is converted into the sentiment of anger itself, by the relation of the idea of anger with the idea (or impression) of our self, which is constantly present to us. The idea of anger gains vividness from the idea of self, which is always vivid to us. Thus the idea of anger is turned into a passion of anger, which can in turn motivate.

Importantly, Hume thinks that our sentiments are constantly changing. As he says, "'Tis difficult for the mind, when actuated by any passion, to confine itself to that passion alone, without any change or variation. Human nature is too inconstant to admit of any such regularity. Changeableness is essential to it." This means that sympathy, which conveys to us the sentiments of others, conveys something that will fade. Sentiments might last beyond the encounter with a person, but the changeability of our sentiments means that the sentiments of others we have acquired through sympathy will also come and go. Even though we acquire the

---

130 T 2.1.11.4, 317. “the idea, or rather impression of ourselves is always intimately present with us”.
131 “All these relations, when united together, convey the impression or consciousness of our own person to the idea of the sentiments or passions of others, and makes us conceive them in the strongest and most lively manner.” T 2.1.11.6, 318.
132 Jennifer Herdt argues that Hume thinks sympathy can be more extended than this: we sometimes sympathize not only with the present sentiment, but also with the future, of another person. See T 2.2.9.13, 385-6 and Herdt, Religion and Faction in Hume’s Moral Philosophy, 46-49. This dramatic extension of the influence of sympathy does not, I believe, change my basic claim that the sentiments acquired through sympathy will be weak motivations to action.
133 T 2.1.4.3, 283.
sentiment of another by association with the vividness the idea of self, the sentiments we acquire through sympathy do not acquire the constancy of the idea of the self. Therefore sympathy will not be able to explain how we are motivated to act on the sentiments of others when we are not directly exposed to them or actively bringing them to mind.

The association of ideas Hume finds in pride, and which is the basis of my account of Humean loyalty, is different from the one we just described for sympathy. As we saw, Hume thinks pride in others is the pleasure we find in the positive attributes and success of an object, person or group associated with the self, while humility is the pain we feel in their negative attributes or failures. Thus pride is caused by an idea of an object associated with the idea of the self, not a sentiment associated with the self: objects that are “in the least ally'd or related to us” (T 2.1.2.5, 278-9). This difference in association has a consequence for the constancy of the motivation. Since our sentiments come and go all the time, when we associate a sentiment with the self it quickly fades. But since objects do not vary constantly, their association with the self does not change constantly. Thus unlike the association that enables sympathy, the association that is the basis for pride offers constancy. There are different ways to achieve this constancy – the relations of contiguity, similarity, and causation can all strengthen the relation between the idea of someone and the idea of the self. But whatever is the way we achieve it, the object associated with the self and the object of self that is extended to include others is (relatively) constant.

134 Although many argue that Hume is skeptical regarding the existence of objects, Stephanie Rocknak has argued convincingly that Hume does offer an explanation for how it is that we hold that objects are constant and continuous. See Rocknak, Imagined Causes.

135 Compare association of ideas in T 1.1.4, 10-13.
Loyalty is a motivation to act on behalf of others that is constant. It often competes with self-interest and often outweighs it – we see this in many instances of loyalty. Thus I conclude that loyalty cannot be a form of sympathy, even a strong sympathy, which is based in associating a sentiment with the self. Rather Humean loyalty is based on the same associations that give rise to pride. Pride in other people is based in the association of other people with the self. This association of objects, which are constant, with a self, which is constant, give rise to a stable motive to act on behalf of those who are part of an extended self.

We can now clearly distinguish sympathy from the motive of loyalty. Sympathy allows a person to feel the sentiments felt by another person, and these sentiments may weakly motivate a person to act on behalf of another. But such actions will be altruist actions, motivated with concern for another only. The association of ideas, on the other hand, enlarges the sense of self of the agent, so that some other individuals or groups become a part of the idea of self. Loyalty, which is the motive to act on behalf of those the self extends to include, strongly motivates a person to act on behalf of a subset of individuals and groups which are part of the self. This motivation is different both in strength and in kind from the motives gained through sympathy.

4. Critique of Hume’s Explanation of the Love of Relations

Before we go on to discuss the virtue of Humean loyalty, there is still a concern that we have overlooked Hume’s account of the love of relations. This seems to be the most natural place to find Hume’s own account of our motive to take up the interests of those who are part of
our extended self, without offering a new motive that Hume did not discuss. Why do we need this new motive that I’ve suggested, if Hume thinks that he has already explained this phenomena?

Most of Hume’s discussion of love (and hate) is not very relevant to our discussion here. This is because Hume characterizes love as finding pleasure in a quality of another person, and hate as pain in a quality of another person, with no reference to the self. Our discussion here has focused on loyalty as a motive to act on behalf of those who are related to and constitute a part of the self. But Hume is aware that we have some form of love, what he calls love of relations, which is not caused by other peoples’ qualities. In the section titled “Of the Love of Relations”, Hume points out that “Whoever is united to us by any connexion is always sure of a share of our love, proportion’d to the connexion, without enquiring into his other qualities.”

This love that we have towards close associates is a problem for Hume’s theory of love, which he explained as caused by the pleasurable qualities of another person. Hume acknowledges that this is not the case, when it comes to the love of relations.

Hume’s solution is to say that although love is usually explained by finding pleasure in the qualities of another person, love of relation and acquaintances is caused differently. Hume observation is that in human beings find their company wanting, and enjoy the change of emotional scenery that is possible with the sharing of the emotions of others. Thus “company is naturally so rejoicing, as presenting the liveliest of all objects, viz. a rational and thinking Being like ourselves, who communicates to us all the actions of the mind… This being admitted, all the rest is easy. For as the company of strangers is agreeable to us for a short time, by inlivening our

---

136 T 2.2.4.2, 352.
thought; so the company of our relations and acquaintance must be peculiarly agreeable, because it has this effect in a greater degree, and is of more *durable* influence.”¹³⁷ We find pleasure in sharing in the passions of others through sympathy, and this is more easily done with those we are familiar with. Thus the pleasure that causes love of relations and close acquaintances is not the pleasure of their qualities, as is usually the cause of love, but rather the pleasure of sharing in their passions.

Why would Hume offer this explanation, which doesn’t fit with his account of love? Hume is avoiding the two options offered by his predecessors: Mandeville argued that love of relations is rooted in self-interest,¹³⁸ while Hutcheson argued that is a form of selfless beneficence.¹³⁹ Hume looks for an answer that will not be self-interested, and yet will be able to explain the special place of these close relations, showing they work differently than the love we have for other people.

I think Hume is right about two things: first, that our love of relations is not dependent on their qualities, and second, that that we enjoy sharing in the passions of others, and that we more easily share in the passions of those related to us. But I think his explanation for why this is, is lacking. Hume claims that the reason why we more easily enter into the passions of those related to us is because we are more familiar with them. “The mind finds a satisfaction and ease in the view of objects, to which it is accustom’d, and naturally prefers them to others,

---

¹³⁷ [2.2.4.4-5, 353.]
¹³⁸ “What we do for our Friends and Kindred, we do partly for our selves: When a Man acts in behalf of Nephews or Nieces, and says they are my Brother’s Children, I do it out of Charity; he deceives you: for if he is capable, it is expected from him, and he does it partly for his own Sake: If he values the Esteem of the World, and is nice as to Honour and Reputation, he is obliged to have a greater Regard to them than for Strangers, or else he must suffer in his Character”, Bernard Mandeville, *An Essay on Charity, and Charity-Schools*, 285-286.
which, tho’, perhaps, in themselves more valuable, are less known to it. By the same quality of the mind we are seduc’d into a good opinion of ourselves, and of all objects, that belong to us.”

But it is hard to see how this is the explanation of the devotion or the motivation we find to improve the situation of our relations. First, this tight connection between relations and ourselves is not limited to or dependent on familiarity. There are no end to stories of family members who meet after being apart for many years, and the association with the self is often not diminished with lack of familiarity. More commonly, people are away from family for a few years, and their emotional tie with family members, and their willingness to do things for them, do not diminish. It seems that the very relation of sister or mother has the effect of making it easier to enter into the sentiments of these individuals, even without familiarity.

Second, familiarity exists with those that we loath, and not only with those that we love. A teenager may be quite familiar with the school bully, and this familiarity will not lead him to love this bully but to hate him. The problem with familiarity is that it only exposes us regularly to those qualities that cause love and hate, and can augment their influence. It does not offer a different way of relating to a person. Thus while Hume acknowledges that love of relations does not depend on pleasure in the qualities of these people, he doesn’t succeed on offering an alternative.

The theory of the expanding of the self that I suggested above does offer an alternative to love of relations that does not depend on, although it may lead, to pleasure in the qualities of those who are related to us. Hume admits that the love of relations is linked with the

\[140\] T 2.2.4.8, 355.
way we relate to ourselves and to “all objects, that belong to us”. Hume’s example is that we come to love the city that we move to, even if at first we find it ugly. But the phenomena here is that we come to associate the city with ourselves, just as our family members are part of ourselves. One way this happens is through familiarity, but sometimes this is achieved through a causal story where our family is part of how we came into the world, with no familiarity whatsoever.

Thus I don’t disagree with Hume that there is a warm fuzzy feeling of love we have for our close relations. I don’t disagree with him that this comes with a stronger beneficence, and with a stronger mutual influence of passions that we find enjoyable in itself. What I disagree with Hume is about the root cause of all this. The root cause is not familiarity, but rather an expansion of the self that we saw in Hume’s account of pride. The very association of another person or city with ourselves leads us to treat them more like the way we treat ourselves, although we still view them as distinct metaphysically, physically and with regard to their will.

Hume might respond and say that this very association with the self is caused by familiarity, by the constant conjunction of ideas between the self and the other individual. Although I’ve accepted for the sake of this account of Humean loyalty that constant conjunction of ideas causes the association of idea that expands the self, I think that we can distinguish between that and the familiarity Hume is discussing in this chapter. As I suggested above, familiarity sometimes does and sometimes does not cause expansion of the self to include others. Expansion of the self is not dependent on familiarity, and can be an association of another person, “my sister”, with my self, with no familiarity with that person at all. A constant
conjunction of ideas can occur with no familiarity of the particular person, and thus the association of ideas that expands the self cannot be reduced to familiarity.

My conclusion is therefore that the love of relations is not, most fundamentally, a familiarity or approval of their quartiles. Although Hume is right that we have some tendency to be “seduc’d into a good opinion of ourselves”, we are also highly critical of ourselves. This is even truer of our relations – we may be more likely to judge them favorably, but their weaknesses and failings also loom in our mind much larger than those of other people. The expansion of the self to include others is what explains the particular characteristics of love of relations that Hume points to: the love that is not dependent on attraction to qualities of the person, which leads to stronger influence of their passions and to the stronger inclination to act on their behalf. I leave open the possibility that if Hume’s account of love is correct, then it is possible that loyalty will lead to the kind of love that Hume is describing. This is because Hume argues in many places that similar passions lead from one to the other. Thus Hume thinks that pride can lead to love of the same object. The same can be said of loyalty and love. But the two passions are distinct, and the root cause of our relation to family members is loyalty rather than familiarity.

5. The Virtue of Loyalty

Like its counterpart, pride, loyalty has often been viewed with suspicion, its particularity causing its virtue to be questioned. Just as others have argued that Hume’s
psychology allows to consider pride a virtue, I argue that it also explains why we judge loyalty to be a virtue. Loyalty benefits others by motivating us to act on behalf of those who are in a special relationship to us. Loyalty also benefits the self, by motivating a person to act on behalf of those whose success will cause the agent a sense of accomplishment and importance.

Hume’s investigation of the virtues leads him to point out the causes of moral approval, which are the shared characteristics of all virtues. Hume concludes that we approve of motives useful or agreeable to self or others. Let us then first consider how loyalty is useful and agreeable to others. Although Hume argues that our sympathy is stronger when it is conveying the sentiments of those close or similar or related to us, we’ve already seen that sympathy can only convey those sentiments that a person has. Loyalty, on the other hand, is caused by an association of a person, along with that person’s needs and interests, with the self. Thus loyalty is more useful to close associations than beneficence. The loyal person is motivated to strengthen and improve the situation of those closely associated with her, regardless of their sentiments. Not only is the assistance useful, the loyal motive is more dependable. This assurance is especially useful to those associated with the loyal person. Not only is it useful, it is also agreeable. Unlike beneficence, which can appear condescending, the actions of the loyal person are agreeable because they are an acknowledgment of a special relationship between the person acting and the person being assisted.

---


142 T 3.3.1.30, 591.

But how is loyalty agreeable or useful to the self? One could suggest then that the ability to prioritize those close to us gives us focus and direction. This view is promoted by Josiah Royce: “Loyalty… tends to unify life, to give it center, fixity, stability.” Such a unification of the self becomes more urgent on Hume’s view of the self as a bundle of impressions, and our ability to sympathize with any human. But although this seems plausible, on Hume’s psychology this goal is achieved by pride. As Amelie Rorty points out, without pride a person will act on the strongest promise of pleasure and will avoid the strongest pain, not only her own but also those of others felt through sympathy. This means that she may act to benefit others, but mostly when it doesn’t cost much to herself. She may prioritize the pain or pleasure of another, but only if she is overwhelmed by their sentiments. Rorty argues that it is pride that saves us from this situation, allowing a person to weigh possible pleasures and pains not only based on their strengths but also according to their importance to herself: “The structure revealed by pride goes beyond the wanton’s push-me-pull-you response to immediate pleasures and pains. Because it gives a differential and independent weight to some pleasures as being closely connected to the self, it goes beyond the projection of the strengths of present pleasures to expectations of the future.” Thus pride is the source of fixity and direction often attributed to loyalty.

Others argue that loyalty is a virtue because it is necessary for maintaining relationships. Kleinig argues that loyalty mitigates egoism and motivates us to uphold and

---

145 T 2.1.4.3, 283.
146 Rorty, “‘Pride Produces the Idea of Self’: Hume on Moral Agency”, 263.
147 Ibid.
strengthen our associations which contribute to the flourishing of the individual.\textsuperscript{148} What is implied by Kleinig’s view is that without loyalty a person would only be motivated by egoism. Ewin is clearer, arguing that without loyalty “groups would not be formed and we should all be isolated asocial atoms.”\textsuperscript{149} If what is implied is that loyalty is the only way in which human beings are motivated non-egoistically, then this clearly does not fit Hume’s psychology. As we saw, on Hume’s psychology human beings are susceptible to the sentiments of others by the mechanism of sympathy. The sympathetic person will not be egoistical on Hume’s view, though she may only be weakly motivated by the sentiments of others.

Oldenquist proposes a third theory, that a person who is not loyal will be alienated. This person will experience bitterness, lack of confidence or guilt that go along with ignoring an association already in place. This does not mean that over time, this association cannot be so much eroded that these effects will dissipate. But long before this happens, a person will sense that he is ignoring part of himself. On this view a person who is not loyal, and does not act to strengthen those associated with him, will not be an atom but an alienated self.\textsuperscript{150}

Humean loyalty helps us understand why it is Oldenquist’s view of the non-loyal person that is most precise. On Hume’s view, a person who is not loyal will not be egotistical – she will be motivated by the sentiments of many people she is exposed to, and these might even motivate her to action at times. What the loyal person lacks is a motivation to prioritize those associated with her self. Notice that the non-loyal person will have others associated with her self, just as other people do. This is because the association of ideas between the idea of others

\textsuperscript{148} Kleinig, \textit{On Loyalty and Loyalties}, 79-84.
\textsuperscript{149} Ewin, “Loyalty and Virtues”, 418-419.
\textsuperscript{150} Oldenquist, “Loyalties”, 187-191.
and the idea of the self, which extends the self to include others, does not depend on the motive of loyalty but rather precedes it. Thus a non-loyal person will not be able to avoid the passions of pride and humility that arise from these associations, finding pleasure in the success of those associated with herself, and finding pain at their failures. But such a person will not be particularly motivated to act on behalf of those who cause her pride or shame, especially when doing so will require combatting the constant motivation of narrow self-interest. Such a person will feel alienation, because she will feel pride and shame but will not be motivated to take action to improve the situation of those she is associated with. This lack of motivation will mean that such a person will be disempowered to shape the well-being of her extended self. Unlike her own egoist well-being, which motivates her to action, the success and failure of those associated with her will have no motivating force for her.

Loyalty is useful to the self because it directs our actions and efforts towards those associated with the self, allowing a person to allocate time and effort in a way that does not ignore the association of ideas that make up the broad self. This allows for a sense of meaning and success in directing one’s efforts of strengthening others towards those who are also associated with the self. This strengthens a person because it allows a person to meet the emotional demands placed on him or her. This loop of effort and success is only possible when we choose to strengthen those associated with the self over time, rather than following the constantly changing objects of sympathy. It creates a loop in which we put our efforts into helping those very same people whose well-being influences us most.
6. Objection: Hume’s Disapproval of Faction

In part 5 I argued that on Hume’s own criteria for virtue, we can explain why we judge loyalty to be a virtue. But did Hume himself think that loyalty was a virtue? There is some room to doubt this, since Hume speaks forcefully against parties and factions, whose members he thinks are motivated by something like loyalty: “an affection to the persons with whom they are united.” I therefore turn to Hume’s account of parties and factions, in order to argue that although Hume thinks some forms of side-taking are problematic, my view of loyalty fits with his account of factions “most reasonable”.

Hume argues that “As much as legislators and founders of states ought to be honoured and respected among men, as much ought the founders of sects and factions to be detested and hated”. Legislators and founders of state give us laws, which have the strongest influence on our virtuous characters. In light of this importance of laws, Hume argues, we can understand why factions are so problematic: “the influence of faction is directly contrary to that of laws. Factions subvert government, render laws impotent, and beget the fiercest animosities among men of the same nation, who ought to give mutual assistance and protection to each other.”¹⁵¹ Hume seems to think that the strong passion of loyalty to faction leads people to prefer the group that they are part of over the laws that they share with the whole nation. Examples for this abound: a party member steals money to run his party’s campaign, or assassinates the candidate of another party, thinking that the laws of the country do not hold when the needs of his own party are at stake.

But it is interesting that Hume does not argue, as some critics of loyalty do, that loyalty perverts our judgment of some rational moral laws. Rather Hume argues that loyalty is harmful because it engenders animosity “among men of the same nation, who ought to give mutual assistance and protection to each other”. Thus Hume is not opposed to any partiality. He is rather stating that our primary loyalty should be to those of our nation with whom we share laws that direct our actions and form our virtuous character, and not with sub-factions among them. Why would this be? On what grounds does Hume think we should prefer our national affection over our party loyalties?

In order to understand this, we need to look further into the kinds of parties Hume discusses. Factions may be divided into personal and real. The personal faction is very common in small societies, where “Every domestic quarrel, there, becomes an affair of state.” Here Hume warns us against the tendency to take sides on every issue. Some disagreements we should not take sides on, even if we care very much about the disputants and are loyal to them. Loyalty does not always require taking a stance on a quarrel that your sister or mother is in, because it might not forward their interest for you to take a side. Moreover, taking a side if you must doesn’t necessarily mean that you enlist others in taking a side as well. There is no need to turn a personal quarrel into a public one, when the consequences of it are not important to the public.

In opposition to personal factions, real factions are those that are fundamentally public. “Real factions may be divided into those from interest, from principle, and from

\[^{152}\text{Ibid, 56.}\]
Of all factions, the first are the most reasonable, and the most excusable.”

Hume’s claim is that the existence of factions, and the passionate loyalty of their members, is most reasonable when there is a conflict of interest between the parties. Thus for example the interests of nobles and commoners, or of soldiers and merchants are really in conflict sometimes, so it is most reasonable that they will “contract on an affection” to those who share their interest and work together to promote their common interest. Hume is most critical of those parties that are formed based on some principle, where the disagreement does not affect the way one lives one life. The only reason why people fight about these is because human beings are “wonderfully fortified by an unanimity of sentiments”, and are “shocked and disturbed by any contrariety”. This leads them to be impatient with disagreement, and want to argue or fight it out. But in these cases Hume thinks that we should just accept that there is a difference of opinion, without viewing those who share our opinion on impractical matters as more on our side and be motivated by a passionate loyalty to prefer them to others.

The third kind of real difference that leads to faction is an argument over who should rule. Here Hume says that the attachment is to the particular individual. Hume points out the strangeness of this phenomena: “it may seem unaccountable, that men should attach themselves so strongly to persons, with whom they are no wise acquainted, whom perhaps they never saw, and from whom they never received, nor can ever hope for any favour… We are apt to think the relation between us and our sovereign very close and intimate. The splendour of majesty and power bestows an importance on the fortunes even of a single person.”

---

153 Ibid, 59.
154 Ibid, 60-61.
155 Ibid, 63.
personal kind of faction is not based on an assessment of the better leadership of the individual, and on the interest of society, but rather only in the great interest we develop in the fortune of this single person. This interest is similar to that we develop in the lives of movie starts, rather than on the interest of our society and its members.

We see in this discussion of the good and bad kinds of faction that Hume does think there is some merit to loyalties based on interest. I think this sheds light on his view of why loyalty to the nation and to its laws should outweigh party loyalties. As we saw in our discussion of justice in the previous chapter, Hume thinks that conventions such as justice are morally approved of because they are a system that is beneficial to all its members. Hume thinks that the laws of a country are beneficial to all, and therefore loyalty to them is a kind of loyalty based on interest and not merely on principle. Although Hume doesn’t say that factions undermine the loyalty we have to the nation, he does say that factions undermine the “mutual affection and assistance” nation members should have to one another. As shall be discussed in the last chapter, the best motive that explains this mutual affection is loyalty.

Hume is so out-spoken regarding faction because he thinks human beings have “a strong propensity to such divisions.” They will divide into group on any minor difference, no matter how silly and not important. Once such divides are created, human beings will hold on to them, because of the affections that are attached to them: “When men are enlisted on opposite sides, they contract an affection to the persons with whom they are united, and an animosity against their antagonists.” But this, I argue, doesn’t point to a problem with loyalty, but with how we utilize it. Loyalty, I suggested, is acting on the interests of those who are part of the

156 Ibid, 58.
extended self. But merely preferring those who are part of the extended self, with no benefit to them and with great harm to society at large, is not loyalty. Participating in an ongoing feud, when both sides would benefit from ceasing this war, is not loyal but harmful to those who are part of the extended self.

How can we tell the difference? It isn’t that easy. Hume himself points out that factions may be created because of real differences of interest, but then the interest ceases and the competition or war continues.\textsuperscript{157} This is because in loyalty we respond to the danger to our object of loyalty, and once a competition or war begins, it could seem that winning in this competition or war is always beating the other side. But this is not always the case. Sometimes a loyal person may be able to bring more benefit to their side by ending the war rather than continuing it. But this kind of insight cannot be gained by someone who is part of a faction that is based in principle or personal war, because these do not change and are not affected by the changing circumstances. This is why Hume prefers the faction of interest, which allows for reevaluation when the situation changes.

Let us take an example, the war between the Montaigne’s and the Capulets, memorialized in Shakespeare’s Romeo and Juliet. The war is portrayed by Shakespeare as one between personal factions: there is nothing to their quarrel other than personal antipathy. This does not change, and Hume claims cannot change, because it pulls more and more people into the two camps, creating more and more personal antipathy. Hume’s claim is that had the war been one of interest – let us suppose that Romeo and Juliet were heirs to competing businesses, then Hume would say that their competition could change when a new competitor entered the

\textsuperscript{157} Ibid, 58.
market, or one developed their business in another direction. Similarly, if their war was based on a difference of social-economic rank, the animosity could dissipate with the change in one family or the other’s economic situation. But if the disagreement were one of principle that is not connected to their own interests, then no change in the situation of either could change the war and it would continue on forever.

**Conclusion**

In this chapter I argued for a novel theory of loyalty as acting on the interests and needs of those who are part of one’s extended self. I first showed that Hume’s theory of the agential self, as viewed in his discussion of the passions, does not have set boundaries and extends to include others. I then argued that the motive to act on behalf of those who are part of the extended self is not pride or self-interest, but rather a distinct motive of loyalty. Finally I showed that this motive is a fourfold virtue on Hume’s account of the virtues. Most notably, it is useful to the self because it allows the individual to devote time and energy to others in a way that strengthens the individual’s extended self.
Chapter 3: Loyalty as the Solution to Hume’s Problem of Obligation

In chapter 1 I argued that according to Hume we feel moral sentiments when we evaluate either the characters of persons or systems of conduct. That chapter was meant to offer a close analysis of Hume’s texts, without critiquing his position. In this chapter I embark on a different project, which is only partially interpretative. I am interested in what could explain moral obligation, on Hume’s moral theory. The question I am interested in is why we experience some moral considera tions as more pressing than others, such that we describe them as binding. It is common to describe moral obligations as actions that we find that we must act upon, even when we do not want to. As I will argue, Hume was greatly interested in this question but his answer is not empirically satisfactory. My aim is not merely to describe this phenomena but to look for a psychological explanation for why we find some moral actions binding, on a Humean picture of our moral psychology.

I aim to answer this question by utilizing Hume’s psychological method, which is an empirical study of the impressions that give rise to our ideas. My study will therefore not draw on the literature that explains obligation by reference to reasons, rationality or phenomenology. Instead, I ask what is the impression, common to instances of moral obligation, which gives rise to the idea of moral obligation,\(^\text{158}\) and what are the causes of this impression. I show that the impression that gives rise to the idea of obligation is the reflective impression of the motive of loyalty. I then show that obligations are actions that fulfill the needs of those people and groups who are part of the extended self.

\(^{158}\) For an explanation of this terminology see footnote 3 in the introduction.
This investigation showing that loyalty is the psychological source of obligation is divided into two chapters. In chapter 3, I utilize Hume’s psychological method to investigate the idea of obligation. I aim to understand what is the impression that gives rise to this idea, what are its causes, and what is the function it plays in our moral life. In order to make this investigation as simple as possible, I focus primarily on examples of obligations in personal relationships. But I do not mean to imply that loyalty is the source of personal obligations alone. In chapter 4 I extend this theory to impersonal obligations. This separation into two chapters is meant to lend simplicity to the argument, by first bringing the psychology into focus in the most obvious cases and then showing how it extends to other less obvious ones. This separation also allows me to more carefully interpret Hume’s discussion in the Treatise, which distinguishes between natural virtues, which give rise to personal obligations, and artificial virtues, which give rise to impersonal obligations.

Although many have written about the place of loyalty in our moral judgment, and some have argued that loyalty is one source among others for obligations, there has not been a view put forth that all moral obligations come from loyalty. My account grapples with two central challenges to this view. One challenge, which I address in chapter 3, is that obligations (at least many, if not all) are agent-neutral. Since loyalty is agent-relative, and obligation is thought to be agent-neutral, it is thought that loyalty cannot explain obligations. I argue that this challenge is caused by a conflation between goodness and obligation. Since moral goodness is agent-neutral, when moral goodness and obligation are conflated it is thought that obligation, too, is agent neutral. I draw on the framework set by W.D. Ross, who distinguishes between the concepts of good and obligatory. I argue that while concepts of moral goodness – virtue and vice,
good and bad – are agent neutral, obligations are agent-relative and thus can have their source in loyalty.

The other challenge to the view that all moral obligations have their source in loyalty is the impersonality of many obligations. It is thought that since obligations – such as promising keeping or returning a loan – are impersonal, they cannot be explained by loyalty that is partial. It is thought that impartial obligations are obligations because they come from impartial rules. I respond to this challenge in the next chapter, where I show that on Hume’s moral psychology impartial obligations arise from conventions. These conventions, although they apply impartially to those who are members of the convention, are partial in that they are adopted by a particular group or society. On this view, I argue, loyalty to the group who adopted the convention is the motive that gives rise to impartial obligations.

Since this study aims to be empirical, it is important to note on the phenomena that I study. Throughout this chapter and the next, I try to use examples of obligation that I and my reader will agree are examples of obligation. At the same time, my position regarding the psychological source of moral obligation has implications for normative ethics, specifically for the list of actions that are moral obligations. As will become clear from my discussion in the next chapter, I believe that our list of moral obligations is shaped by our loyalties, and therefore I do not think that it is possible that scholars from different religions, traditions and societies will have the exact same normative views. I therefore do not expect my reader to share my normative view. I merely hope that there is sufficient overlap in our normative theories, so that we have a set of agreed-upon obligations we can analyze, in order to bring into view the common psychology that I believe shapes any normative view.
This chapter will proceed as follows: in part 1 I describe the problem of obligation, as Hume saw it, and present the best interpretation of Hume’s account of moral obligation as arising from the moral sentiments. I critique this view, and conclude that we need to conduct a separate psychological investigation of obligation. In part 2 of this chapter I apply Hume’s experimental method to show that obligations are explained by the causes of loyalty, and thus that obligations are actions that fulfill the needs of those who are part of one’s extended self. In part 3 of this chapter I argue that although obligation has its psychological source in a motive different from moral approval and disapproval, a Humean can remain a judgment internalist by rethinking Hume’s account of moral judgment. I offer a new theory of moral judgment of moral obligation as the calibration of the moral sentiments with loyalty.

1. Critique of Hume’s Judgment Internalism

In this part I first present Hume’s problem of obligation which I aim to answer: what is the connection between moral judgment and the will to act morally. I then proceed to present what I take to be the best interpretation of Hume’s view of obligation as a judgment internalist about obligation. On this view, obligation arises when we direct the moral sentiments at ourselves and we morally disapprove of our own character. I then critique Hume’s solution: I argue that this solution does not take seriously enough Hume’s claim that in order to understand obligation we must look at the influences on our will, because moral sentiments and obligation have different influences on the will. We can distinguish moral approval and disapproval, which apply in a wide array of instances that don’t influence the will to action, from obligation which
appears in instances where the will is influenced to act. Since obligation dictates action, while moral approval and disapproval do not, we will need to investigate obligation as a distinct phenomenon in order to more fully understand the connection between moral judgment and moral action.

1.1 Hume’s Problem of Obligation

In this section I present Hume’s problem of moral obligation. Hume asks what psychologically explains our moral obligation. He thinks that we find an answer to this problem by studying the connection between moral distinctions and our will to act morally.

In order to understand Hume’s problem of obligation, we must view it in light of his explanatory project. There was a common way of reading Hume on which Hume was understood to be a thorough skeptic. Such a reading took Hume to be arguing that many of the important concepts we use do not refer to anything. This reading of Hume as an error theorist has mostly been discredited in the last few decades. Scholars have shown that Hume is not bent on explaining why we are wrong to use such concepts as objects, identity or causation. As Don Garrett and Stephanie Rocknak, among others, have argued, his aim is rather psychological, to show the psychological mechanisms and patterns that lead us to use such terms.\(^{159}\)

While we can debate what Hume thought there was any source in our impressions for the idea of substance,\(^{160}\) it is quite clear that Hume found extreme skepticism with regard to

---

\(^{159}\) Garrett, *Cognition and Commitment*, 10, 202; Rocknak, *Imagined Causes*, 3, 6. This is not the case for all concepts – Hume does think that some concepts are nonsensical, such as substance. But his overall project is not skeptical. See for example the positive account of causation that Garrett and Rocknak think he puts forth. Garrett, *Cognition and Commitment*, chapter 5; Rocknak, *Imagined Causes*, chapter 2.

\(^{160}\) T 1.4.5.5, 233; see discussion in Hazony and Schliesser, “Newton and Hume”, 7.
moral ideas impossible for human beings. In the opening to his *An Enquiry Concerning the Principles of Morals* Hume writes that it is useless to talk to someone who pretends that moral concepts do not refer to something real: “Those who have denied the reality of moral distinctions, may be ranked among the disingenuous disputants; nor is it conceivable, that any human creature could ever seriously believe, that all characters and actions were alike entitled to the affection and regard of every one.”\(^{161}\) Thus in this work I will follow those who interpret Hume as a cognitivist, who holds that moral concepts do refer to facts, and that we can make statements regarding moral facts that are true or false.\(^{162}\)

One passage that is read as a skeptical claim about morals is Hume’s famous is/ought paragraph. On the non-cognitivist reading, Hume here argues that we cannot derive statements about moral obligations from statements about facts, because moral concepts are categorically different from natural ones.\(^{163}\) The famous passage reads as follows:

> In every system of morality, I have hitherto met with… I am surpriz’d to find, that instead of the copulation of propositions, *is*, and *is not*, I meet with no proposition that is not connected with an *ought*, or an *ought not*. This change is imperceptible; but is, however, of the last consequence. For as this *ought* or, *ought not*, expresses some new relation or affirmation, ’tis necessary it shou’d be observ’d and explain’d; and at the same time that a reason shou’d be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it.\(^{164}\)

\(^{161}\) M 1.2, 169-170.


\(^{164}\) T 3.1.1.27, 469-470.
Although Hume says here that “ought” propositions are “entirely different” from the propositions from which they are often supposedly inferred, he doesn’t say that they are different because propositions regarding obligations are not referring to facts. I follow the interpretation set forth by W.D. Falk and Rachel Cohon, on which Hume is making no categorical distinction between facts and values and is not claiming that we cannot derive normative claims from factual claims.\footnote{Falk, “Hume on Is and Ought”, Canadian Journal of Philosophy, 6 (3), 1976, 359-378, especially 369-370. Falk argues that Hume thought moral facts are not reduced to other facts, but this doesn’t mean they do not supervene on matters of fact; Cohon, Hume’s Morality, 11-12, 25-28, and especially 93-95.} Indeed, Hume often derives conclusions regarding the virtue or vice of some character trait from facts of our approval or disapproval for certain motives. Thus in the above quote Hume is not promoting a non-cognitivist view. Rather he is arguing that those facts which are often inferred from cannot on their own explain the moral evaluation that proceeds from them.\footnote{Cohon, Hume’s Morality, 106.} Hume is not arguing that we cannot infer obligations from facts, he is claiming that the facts cited in such inferences are of the wrong kind. This is why Hume thinks that “every system of philosophy I have hitherto met with” has not explained how we go from describing what is good and bad to explaining what we must do, and he raises the challenge: “this ought, or ought not… shou’d be observ’d and explain’d”.\footnote{Although the quote, as it is presented, makes it sound as if this is a question only about propositions, in context we see that Hume speaks not only about propositions but about moral reasoning more generally. See T 3.1.1.27, 469-470.}

As we saw in chapter 1, Hume’s view of ideas is empirical: ideas are copied from impressions. Hume thinks that an explanation of an idea is given by distinguishing an impression that is the source of the idea, and then pointing out the causes of this impression. Thus if Hume
were to seek an explanation of the idea of obligation, Hume would want to discern what is the impression that gives rise to the idea of obligation\textsuperscript{168} and then find the causes of this impression.

Hume thinks that there are two kinds of impressions, impressions of sense, that come from the five senses, and reflective impressions, what we today sometimes call feelings, which are original responses to felt or remembered pains or pleasures.\textsuperscript{169} To which kind of impression does Hume think we should trace our idea of obligation?

In his chapter arguing against moral distinctions being derived from reason Hume gives us a clear view of what he takes to be the missing component of the moral theories he objects to, and thus what is the main criterion for the impression that is the source of the idea of moral obligation:

’Tis one thing to know virtue, and another to conform the will to it. In order, therefore, to prove, that the measures of right and wrong are eternal laws, \textit{obligatory} on every rational mind, ’tis not sufficient to shew the relations upon which they are founded: We must also point out the connexion betwixt the relation and the will; and must prove that this connexion is so necessary, that in every well-dispos’d mind, it must take place and have its influence; tho’ the difference betwixt these minds be in other respects immense and infinite.\textsuperscript{170}

Here Hume argues that if we are to explain obligation, “that the measures of right and wrong are eternal laws, \textit{obligatory} on every rational mind”, we must not only “shew the relations upon which they are founded”, but “point out the connexion betwixt the relation and the will”. Hume says that in order to explain obligation we must find its motivating source. Thus if we show that obligations come from God’s commands, or some internal goodness, or from mind-independent facts of goodness, without showing that these influence “every well-dispos’d mind” to action,

\footnotesize
\textsuperscript{168} T 1.1.2.1/7-8; Radcliffe, \textit{Hume, Passion and Action}, 17-18.
\textsuperscript{169} T 3.1.1.22/465-466, emphasis in the original.
\textsuperscript{170} T 3.1.1.22/465-466, emphasis in the original.
then we have not shown that they obligate. According to Hume, in order to explain what
obligation is it is not sufficient to describe and explain the difference between good and bad. We
must explain the relation between the source of moral distinctions of good and bad and the will
of human beings to act morally. In order to explain why we “ought” to act morally, or why we
are obligated to do so, we must find a motivating source for these ideas.

Hume here is taking the position that moral obligation must be, according to the term
among the British Moralists writing in the 17\textsuperscript{th} and 18\textsuperscript{th} century, to search for the explanation of
obligation in a source internal to human psychology and to the psychology of the obligated
individual. Although they gave different accounts of what this internal source is, these moralists
agreed that there is a motivating source that is internal to the agent that is the source of
obligation.\footnote{Stephen Darwall, \textit{The British Moralists and the Internal Ought 1640-1740}, Cambridge University Press, New
York, 1995, 12-14.} Hume’s framing of the question of obligation fits with this description of the
problem as discussed in the context he was writing in. He raises the question, what is obligation,
and tells us that the way to answer this question is to look at the connection between moral
distinctions and the influences on our will.

But unlike other moralists of his time, Hume thought that only the passions motivate
our will to action. Hume famously argued that moral distinctions cannot have their source in
reason alone. Hume’s argument was based on his own theory of reason: reason compares and
contrasts ideas, and thus does not, on its own, motivate action.\footnote{Garrett, \textit{Cognition and Commitment}, 93; Cohon, \textit{Hume’s Morality}, chapter 3; Radcliffe, \textit{Hume, Passion and Action}, chapter 2.} Reflective impressions are the
ones that influence our will and motivate action. Motivation comes from reflective impressions, or passions, which motivate action because they are forms of pain and pleasure.\textsuperscript{174}

Thus we can sum up that Hume’s question is what connects our moral judgment of good and bad to our will to act morally. This question will be answered, according to Hume, if we look at our motivations to act morally. The motivations to act, according to Hume, are passions. Thus the question is what passion leads us to act morally, and to fulfill our obligation to act according to our moral judgment. I call this Hume’s problem of moral obligation.

1.2 Hume’s Judgment-Internalist Solution to the Problem of Obligation

In this section I present what I take to be the best interpretation of Hume’s answer to the problem of obligation. Hume’s answer is that moral sentiments, which are motivating because they are a kind of passion, are the source of moral ideas utilized in moral judgment. His answer to the problem of obligation is that moral sentiments are the source of our distinction between vice and virtue, and that these sentiments on their own can influence our will to act morally. This is the view I critique in the next section.

Elizabeth Radcliffe offers the best explanation of Hume’s view of obligation in her “How does the Humean Sense of Duty Motivate?”\textsuperscript{175} Radcliffe argues that a close reading of Hume’s texts gives an account of how moral sentiments themselves give rise to obligations. She

\textsuperscript{174} T 2.3.3, 413-418, “Of the Influencing Motives of the Will.”
calls this judgment internalism – on this view Hume thinks that the moral sentiments which are the source of moral judgments themselves give rise to obligations.\textsuperscript{176} This can be distinguished from agent internalism, the view on which the agent has other motives, other than the ones that give rise to moral judgment, which motivate the agent to act according to his moral judgment.\textsuperscript{177} For example, Hume’s view can be distinguished from Francis Hutcheson’s, because for Hutcheson moral judgments are made by a moral sense, while moral action is motivated by beneficence.\textsuperscript{178}

Radcliffe reads Hume as holding the view that obligations arise when we turn our moral approval or disapproval onto ourselves. An example of how obligations arise, according to Hume, can be found in the following passage (which is a major source for Radcliffe’s reading):

> 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 the 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.\textsuperscript{179}

In this passage Hume gives us an example of how obligation arises. A person who approves of parental dedication and disapproves of a parent who lacks it, disapproves of himself when he finds such a lack in himself. This moral disapproval (which is a particular kind of pain) causes a

\begin{footnotes}
\item[176] Radcliffe, “How does the Humean Sense of Duty Motivate?”, 389. Radcliffe points out that in the past it was assumed that Hume was a judgment internalist, 385. She points to the examples of Philippa Foot, “Hume on Moral Judgment”, 79; Mackie, \textit{Hume’s Moral Theory}, 52-53. Radcliffe points out that one could be an \textit{a priori} judgment internalist or an \textit{a posteriori} judgment internalist. Hume makes claims regarding the psychology of human beings, not the necessary character of moral judgment, and thus is a judgment internalist \textit{a posteriori}, see Radcliffe, “How does the Humean Sense of Duty Motivate?”, 392-393.
\item[177] Radcliffe opposes the view of Brown and Darwall that for Hume the motivation of obligations comes from desire for pleasure. See Radcliffe, “How does the Humean Sense of Duty Motivate?”, 394-403.
\item[178] Darwall, \textit{British Moralists and the Internal ‘Ought’}, chapter 8.
\item[179] T 3.2.1.8, 479. Annette Baier thinks that Hume comes to hold the view that only the artificial virtues dictate obligations, see \textit{The Cautious Jealous Virtue: Hume on Justice}, Harvard University Press, 2010, 45. This leads her to struggle with this passage. She distinguishes between Hume’s use of duty and obligation in order to claim that this is not an instance of obligation on Hume’s view.
\end{footnotes}
motivation to act according to the virtue, which in time leads the person to acquire the virtue. This motive to acquire a virtue that one does not have is the source of the idea of obligation, and thus its source is the moral sentiment itself.

Radcliffe’s main argument for this interpretation is that this answer regarding obligation is central to Hume’s argument against the rationalists. Unlike rationalists who sought to find the internalist motivation in human reason, Hume thought that reason was inert, meaning it does not motivate in itself. This is why Hume argues that the moral distinctions, which are motivating, are derived from the sentiments of moral approval and disapproval, which are passions and therefore are of the category of mental states that influence the will and motivate action. Thus the reason why the “connexion is so necessary” between moral judgment and the will is because the source of moral judgment is sentimental and sentiments themselves motivate.

On Radcliffe’s reading, what explains moral obligation is the sentiment of moral approval. This view is opposed by those who think that Hume held that only some of our moral approvals give rise to obligations. Such a view is proposed by Annette Baier, who argues that Hume “tends to reserve the term [obligation] for what the artificial virtues demand.”180 Baier argues that Hume wanted to distinguish moral obligation from the ideas of virtue and vice, since he thought we should distinguish behaviors that are required from behaviors that are encouraged.181 She argues that “In Hume’s moral theory as a whole, deontology is circumscribed and subordinated to the main account of morality as the cultivation and welcoming of virtues,

180 Annette Baier, *The Cautious Jealous Virtue: Hume on Justice*, Harvard University Press, Cambridge, 2010, 45. Bernard Wand also argues that Hume thinks there are only obligations in the case of the artificial virtues He takes obligation to mean acting out of “regard for the moral worth of the action”, Wand, 157, which he thinks is a motive that is only developed in response to a convention (similar to Garrett’s artificial motive discussed in chapter 1).
both natural virtues and artificial virtues.” Baier is right that it is striking that the term obligation is only used in the discussion of the artificial virtues. But she admits that this reading is difficult, given the passage quoted above from Hume’s discussion of the natural virtues, in which he refers to duties.\textsuperscript{183}

I think that Radcliffe is correct in her reading of Hume, because Hume often uses the terms duty or obligation to denote that something is morally approved of. Thus for example Hume asks “But may not the sense of morality or duty produce an action, without any other motive?”\textsuperscript{184} In posing the question, Hume uses “sense of morality” and “duty” interchangeably. Another central passage where Hume uses obligation interchangeably with moral disapproval is the following:

All morality depends upon our sentiment; and when any action, or quality of mind, pleases us \textit{after a certain manner}, we say it is virtuous; and when the neglect, or non-performance of it, displeases us \textit{after à like manner}, we say that we lie under an obligation to perform it.\textsuperscript{185}

Hume’s use of moral obligation in this way is not unusual for his time. Francis Hutcheson reports two uses of the term obligation by his contemporaries: First, “\textit{That the Action is necessary to obtain Happiness to the Agent, or to avoid Misery}”, and second, “\textit{That every Spectator, or he himself upon Reflection, must approve his Action, and disapprove his omitting}”.

\begin{itemize}
\item \textsuperscript{182} Ibid, 136.
\item \textsuperscript{183} In order to solve this problem she suggests that Hume distinguishes between duties and obligations. I think this is supported by the following passage from Hume’s essay “On the Social Contract”: “All \textit{moral} duties may be divided into two kinds. The \textit{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… The \textit{second} kind of moral duties are such as are not supported by any original instinct of nature, but are performed entirely from a sense of obligation.” David Hume, \textit{Essays Moral, Political and Literary}, (Edinburgh, 1741), ed. with foreword, notes and glossary by Eugene F. Miller, Liberty Classics, Indianapolis, 1978, 479-480. But if this is the case, then there is still such a thing as duty, in all virtues, and our discussion of moral obligation could overlap with Hume’s use of duty rather than only obligation.
\item \textsuperscript{184} T 3.2.1.8, 479
\item \textsuperscript{185} T 3.2.5.4, 517.
\end{itemize}
it, if he considers fully all its Circumstances”.

We find that Hume too uses the term obligation in these two ways. He distinguishes between natural obligation, which comes from the motive of self-interest, and moral obligation, which comes from the motives of moral sentiments: “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.”

1.3 A Critique of Hume’s Solution

As we saw, Hume’s solution to the problem of obligation is that the moral sentiments themselves connect between moral judgments and our will. In this section I critique Hume’s solution. I argue that his view of obligations as arising from moral approval and disapproval is empirically inadequate for two reasons: one, it entails too many obligations, and two, it fails to acknowledge the role of obligation in motivating us to act morally in face of strong opposing motives. Since Hume’s view of moral sentiments is that they are motivationally weak, I argue that they cannot explain obligation which motivates us to act morally in face of strong opposing motives.

The first problem with Hume’s use of obligation is that it leads us to define too many obligations. Hume’s view of moral approval and disapproval gives rise to an unlimited number of virtues and vices. If we were to take Hume’s claim that all moral approval and disapproval give rise to moral obligation seriously, then there would be endless obligations. It would mean

---


187 T 3.2.2.23, 498.
that a person must view themselves in light of every virtue, and be obligated to fix any vice they have. We can concede that hating one’s self for lacking the natural motive to care for one’s children can motivate one to act in a way that would lead one to acquire this motive. But on this view, if a person hates herself for not caring for her children, for not having beneficence towards friends who ask for her help, and for not being courageous in facing belittling colleagues at work, then she would be obligated to acquire all of these virtues.

If obligations did not have this limiting factor, and we were equally motivated to act upon every moral approval we had, we would actually be paralyzed by the endless and conflicting motivating moral approvals. If we thought that we ought to act upon every moral approval and disapproval, then we would be constantly torn between which moral approval and disapproval to act upon. The father who hates himself for not caring for his children, while he is home, will then go to work and hate himself for his tardiness, and then go to the gym and hate himself for his laziness. These moral disapprovals wouldn’t tell him how to allocate his time between home, work, and the gym. If he considered each moral evaluation on its own, it would pull him in the direction of spending unlimited time acquiring a particular virtue. But if he considered all these moral approvals and disapprovals at the same time, he would be paralyzed because no moral approval will lead him to prioritize one action over another.\footnote{188}{I address the possibility that stronger moral approval is the source of obligations in the next section.}

The list of ways in which we could improve in light of our moral approval and disapproval goes on and on, but we do not think of obligations as endless in this way. If a person did judge himself in light of every moral standard, this person would hold himself accountable to
act to acquire every virtue.\textsuperscript{189} But in our common sense morality we do not hold ourselves obligated to act in every way that we consider virtuous. We think that some virtuous actions are supererogatory. We use obligation as a limiting factor, which tells us which good actions we must take much more seriously, and view as more pressing, than other moral actions. Although we don’t view obligations as absolute, we consider obligations more pressing, and thus we are not required to act upon all moral evaluations.

The other problem with Hume’s view, on which obligations arise from moral approval and disapproval, is that on Hume’s own account these sentiments are motivationally weak, and cannot overcome self-interest. Empirically we find that many instances in which we are most aware of acting on our obligation are instances where we are motivated to do so against our own self-interest. Let us first understand why the motive that gives rise to the idea of obligation must be a strong motive, and then why Hume thinks that the moral sentiments have a weak influence on our will to act. These together will explain why the moral sentiments are inadequate in explaining moral obligation.

Hume does not discuss the strength of the motive of obligation. He thinks that we can sufficiently explain obligation by pointing to a motive that influences the will. But this picture insufficiency accounts for our experience of acting out of obligation. There are many instances in which we have strong motives to act viciously, in opposition to our moral approval and disapproval, and yet we act virtuously. We need an explanation of what strong motive allows us to act morally in such cases. Hume does not acknowledge this role of obligation. Radcliffe seems

\textsuperscript{189} This is analogous to the criticism of utilitarian views, which seem to hold up an impossible bar on which people have to constantly calculate in what way they can maximize utility.
to think that the strength of the motive is irrelevant to whether or not it generates obligations. But it seems that we become aware of obligation precisely when we have strong motives to act viciously, and we act morally out of obligation.

We can take Hume’s example of the father who hates himself for not caring for his children as an example. Let us concede that a father can hate himself for not caring for his children, meaning he can hate himself for lacking the natural motive, which is approved from a common point of view. Let us also concede that this self-hate motivates, so that in a vacuum, this sentiment could cause the father to act in a way that will help him acquire the motive of care for his children. My contention is that such a father is acting out of moral approval and disapproval, but that this doesn’t fully describe what happens in cases of acting out of obligation. The question of what is obligation, I propose, does not come up in the situation that Hume describes, but rather in cases of conflict in which prioritization is necessary. Let’s say the father does not care for his only son, because this only son repeatedly belittles him, or is a murderer. It is in such difficult situations that we can see that the father acts out of an obligation to care for his son, because doing so is difficult and he has a strong motive of self-interest or moral disapproval of his son that motivate him to not care for his son. My contention is that the motive that gives rise to the idea of obligation is most easily discerned in such instances.

On Hume’s own account, moral approval and disapproval cannot be this strong motive which outweigh self-interest. Hume mentions in several places that moral approval, which is acquired through sympathy with the people who are associated with the person being

---

190 See “My general concern has been with whether there is a way in which the moral sense directly produces a motive to morality… not with whether it produces the strongest among competing motives.” Radcliffe, “How does the Humean Sense of Duty Motivate?”, 404.
appraised, is quite weak. He is even more specific – he says that sentiments acquired through sympathy often influences our judgment of another person’s character and the way we will describe him, but often do not change our actions.

My sympathy with another may give me the sentiment of pain and disapprobation, when any object is presented, that has a tendency to give him uneasiness; tho’ I may not be willing to sacrifice any thing of my own interest, or cross any of my passions, for his satisfaction.\textsuperscript{191}

Hume’s claims here are somewhat confusing, since it is not clear why any certain motive should be strong or weak in general. Perhaps the moral sentiments are weak, but could be strengthened in some way? We could propose the solution that society or education strengthen the influence of the moral sentiments,\textsuperscript{192} and thus the moral sentiments are able to motivate us in situations where it is difficult to act virtuously. But Hume doesn’t limit this claim to non-cultivated people. Why does Hume make a general claim that moral sentiments have a weak influence on the will?

I suggest that moral sentiments are weak on Hume’s psychology because, on his account, moral sentiments are generated through the mechanism of sympathy. Sympathy is Hume’s mechanism by which we become susceptible to the sentiments of others. The way we come to morally evaluate the character of another person is by sympathizing with the people who are influenced by this person. In order to judge this person’s influence on that close circle of associates, rather than from one’s own self-interest, one must sympathize with those in the close circle. Thus Hume observes that

This latter principle of sympathy is too weak to controul our passions, but has sufficient force to influence our taste, and gives us the sentiments of approbation or blame.\textsuperscript{193}

\textsuperscript{191} T 3.3.1.23, 586-7.
\textsuperscript{192} This is discussed by Radcliffe in “Hume on Motivating Sentiments, the General Point of View, and the Inculcation of “Morality””, Hume Studies, 1994, 20 (1):37-58.
\textsuperscript{193} T 3.2.2.24, 500.
I propose that the weakness of the moral sentiments is not a bug, but a feature, of their function. Hume is not pointing to this weakness as something to be overcome.\textsuperscript{194} The moral sentiments don’t arise in order to motivate us to act morally, but rather allow us to distinguish those people and systems of conduct we approve of from those we disapprove of. The function of such distinctions is much broader than pointing out how we should act, although they have implications for how we should act. When we morally approve of a person who lived centuries ago, or of a system of conduct in another country, this may not lead us to any action. This allows us to converse with others and open options for giving advice, enacting conventions that will encourage such action, or coordinate actions with others. Thus the weakness of moral sentiments is a feature of their function, to enable conversation with others about virtues and vices.

If we judge that being beneficent or courageous are virtuous motives, this will only weakly motivate us to act in these ways.\textsuperscript{195} This means that if we have a strong opposing motive, this moral sentiment on its own will not motivate us to act accordingly. This weakness is not overcome, since if we strengthened the influence of moral sentiments they would not be able to fulfill their function as a way to communicate with others about a large spectrum of instances. But if Hume is right that moral approval and disapproval are a weak motivation to action, he

\textsuperscript{194} I here disagree with Bernard Wand, who suggest that Hume’s motivation of obligation is moral approval or disapproval strengthened by social influence and education. See Bernard Wand, “The Originality of Hume’s Theory of Moral Obligation”, 162.

\textsuperscript{195} Radcliffe also points this out late in her argument, see Radcliffe, “The Humean Sense of Duty”, 405. At first it seems that she thinks this is only true in the case of the natural virtues, but in footnote 44 she acknowledges that this is true in the case of the artificial virtues as well.
therefore is mistaken to speak as though these moral sentiments themselves give rise to obligations that strongly motivate action.\textsuperscript{196}

I suggest that when we talk about obligation, we mean precisely what is the strong motive to act virtuously that kicks in when moral approval is “too weak to control our passions”. Thus it is my contention that in order to explain obligation, on a Humean picture of our moral psychology, we must look not to the weakly motivating moral sentiments of approval and disapproval, but rather to another motivation that often accompanies these moral sentiments and is a strong motive to action.

My claim that obligation must be explained by a strong motive to action should not be understood as a claim about how easy or difficult it is to fulfill obligations. Obligations are often quite difficult to fulfill. What I am arguing here is that precisely because obligations are often difficult to fulfill, it cannot be the case that a reflective impression that is only a weak motive to action can explain how obligation works. The role obligation has of prioritizing among our moral evaluations doesn’t ensure that acting morally will be easy. It merely directs our focus to some moral judgments rather than others, so that acting morally will be possible. Acting upon all moral evaluations is not possible, but choosing which moral evaluations will be acted upon makes those chosen into goals we can strive to achieve, whether they are easy or difficult.

This is not meant to imply that explaining obligation will do away with all conflict: our obligations themselves can and do often come into conflict. My contention is merely that we distinguish between conflicts among obligations and conflicts between obligations and mere

\textsuperscript{196} Wand points out that the motive of obligation must be a motive that in fact motivates moral action when it is difficult to act morally. See Wand, “The Originality of Hume’s Theory of Moral Obligation”, 155, 162.
moral approval or disapproval. Overall, we consider obligations to take precedent over other actions we approve or disapprove of, and thus they cannot have the exact same explanation.

To sum up, there are two problems with Hume’s view of obligation as arising from moral approval and disapproval. First, this theory entails too many obligations, which means that obligations cannot fulfill their role in prioritizing among possible paths and directing action. Second, if moral obligation must be a strong motive to action, but Hume is right that we are only weakly motivated by moral approval and disapproval, then moral sentiments are not the right motives to explain obligation which is most obvious (though it may not be limited to) instances where acting morally is hard.

I thus suggest we ask the question in this way: Let us assume that Hume is right, that obligation must be explained by a motive, and that motivations to action are reflective impressions (passions and sentiments). Let us also agree with Hume, that moral approval is motivating, but is a weak motive to action. Given this situation, what is the motive to action in instances where acting virtuously is hard, which gives rise to the idea of obligation? Here Hume does not press his questioning far enough. Although he explains why moral distinctions must have an influence on every mind, since they come from sentiments that motivate, he doesn’t pay attention to the unique influence of obligation, and thus does not give an adequate explanation of it.

1.4 A Motive for Obligation

My above criticism of Hume’s position was that moral sentiments cannot explain obligation, because obligation has the role of prioritizing among our moral evaluations. A
response to this challenge, on behalf of the Humean judgment internalist view we saw above, could be that we prioritize among our moral evaluations by the strength of these evaluations. Thus one could argue that the Humean view of obligation should be that obligations are those acts we most approve or disapprove of. In this section I respond to this suggestion by showing that moral approval and obligation sometimes come apart, and thus we cannot explain obligation in terms of moral approval.

W. D. Ross criticized utilitarianism for holding that right actions are those that maximize utility. Utilitarians fall to this mistake because they have only one moral concept, utility, and thus they try to explain the right in terms of this one concept. Ross argued that the right comes apart from the maximization of utility. He pointed out that, for example, when we judge that we are obligated to keep a promise we make this judgment independently of any calculus of utility. Thus Ross argued that maximization of utility cannot explain what is right. Ross argued that we must distinguish two concepts applied in our moral judgment, the good and the right, which cannot be reduced one to the other.

The suggestion raised above, that the Humean view of obligation should be that obligations are those actions we most approve or most disapprove of, would fall to the same criticism Ross directed towards utilitarianism. This view would depict moral obligation as being derived from the same source as moral goodness. Thus it would not recognize that the good and the right are distinct concepts that cannot be reduced to one another. In the following I

---

198 Ibid, 17. This basic truth is not undermined by the exception cases on which we forgo keeping a promise in order to avoid terrible consequences. See ibid, 18.
199 Ross mentions that right and moral obligation are equivalent, but he finds moral obligation too cumbersome. He is also wary of the possibility of confusing obligatory with compulsory. See ibid, 4.
substantiate the distinction between morally approved and obligatory, not by conceptual analysis but by examples. I show that our moral approval and disapproval come apart from obligation at times. This means that on a Humean picture, we must distinguish between the source of the concept of virtue and vice and the source of the concept of obligation.

First, we can see that moral approval and disapproval come apart from obligation in instances where fulfilling an obligation requires an act that we do not morally approve of. The classic case of Socrates in Athens is a case in point. Socrates insists on fulfilling his obligation to his city by carrying out his own execution, although he thinks that he is innocent and the Athenian jury judged him unjustly. It is clear that Socrates does not morally approve of his execution, since he believes he is innocent. So he thinks that his obligation to carry out the decision of his jury requires that he act immorally and kill an innocent person, namely himself. Even if we think that Socrates should not enact the verdict, and should instead flee, this doesn’t mean that we think he has no obligation to follow the verdicts of his city. This is because while a person may forgo fulfilling an obligation to avoid great evil, this doesn’t erase the obligation.200

One could of course suggest that Socrates morally approves of fulfilling civil obligations, and thus that he judges it morally virtuous to act according to his obligation. This may be the case, but in order for us to say that fulfilling civil obligations is moral we would need an explanation for what obligations are. We first need to explain what acting on one’s obligations is, before we can say that we morally approve of fulfilling our obligations, and my claim is that explaining what an obligation is does not first rest on the view of what is moral.

---

200 On Ross’s view, we have *prima facie* duties are considerations to be taken into account when deciding our single, all things considered, obligation in a certain situation. See Ross, *Right and the Good*, 20. I see no source in Hume for this idea of *prima facie* duties. Pluralism with regard to obligations would fit best with Hume’s moral pluralism. For a discussion of Hume’s moral pluralism, see Michael Gill, *Humean Pluralism*, Oxford University Press, 2014.
The case of Socrates is an extreme one, because it shows that we can morally disapprove of an action that is obligatory. Other instances that show that moral approval and obligation come apart are not so extreme. First, there are many instances in which we approve of an action that we consider supererogatory – donating a kidney to a stranger, for example. In such instances moral approval clearly does not entail an obligation. Second, there are many instances in which we morally approve of two actions, but we only consider one of them obligatory. For example, a doctoral student pressed to complete her dissertation may morally approve of missing a few hours of work to visit a sick friend, but she will consider it obligatory to put aside her work to care for her sick children. Although both actions are virtuous, only one of them is obligatory.

Another place we see that obligation and moral approval come apart is in the lingering sense of obligation that is unfulfilled. We can detect an unfulfilled obligation, even when it is unfulfilled because of moral disapproval, because avoiding the action that is disapproved of does not get rid of the lingering compunction about the unfulfilled obligation. Thus we can imagine that a sister would morally disapprove of supporting a sister who is an alcoholic, because she does not want to encourage her sister’s harmful addiction. But making this choice will be accompanied with a constant regret, a nagging feeling that an obligation has not been fulfilled, because the alcoholic sister is being neglected rather than cared for. This regret does not arise because the sister does not approve of her own actions, but because moral approval does not nullify one’s obligations.

---

201 Ross points out that we often feel compunction when we don’t fulfill a prima facie obligation, even when it is outweighed by another prima facie obligation. See Right and the Good, 28.
To conclude this part, I do not take Hume to have offered a satisfactory theory of moral obligation. Hume is not skeptical regarding moral obligation, but he does not give a satisfying account of moral obligation. This is because he views obligation as a by-product of moral approval and disapproval, and thus does not investigate the concept of obligation as a concept distinct from the good. Because Hume held that the moral sentiments motivate in a weak way, and we often need a strong motivation to act morally in light of conflicts with self-interest, Hume did not sufficiently explain the connection between our moral judgments and our will. I propose that we conduct a Humean enquiry of obligation, and find the motive which, together with the moral sentiments, motivates moral action and is the source of the idea of obligation.

2. Loyalty: The Impression that Gives Rise to the Idea of Obligation

In this part I first outline Hume’s method, and then apply it to the idea of obligation. In show that in order to offer a Humean explanation of obligation, we must describe incidents of obligation, compare them to find the impression that is common to them which allows us to group them together under a single concept, and then find the causes of this impression which explain the idea. I first show that considering instances of obligation offers us a number of criteria for the impression that gives rise to the idea of obligation. I then propose that the motive of loyalty I introduced in chapter 2 fulfills these criteria, and is the impression of obligation. My conclusion is that the impression that gives rise to the idea of obligation is the reflective impression of loyalty, and therefore that obligation is explained by the causes of loyalty: the
needs of those who are part of the extended self. On this view, obligations are actions that fulfill the needs of those who are part of one’s extended self.

2.1 Hume’s Experimental Method

Hume divides all of our mental states (perceptions) into impressions and ideas. All ideas, he argues, are derived from impressions either by copying or representations. Importantly, Hume thinks that there are two kinds of impressions. Impressions of sense seem to come from outside of us (these are the ones we talk about as impressions of the material world) and are connected to one of the five senses, while reflective impressions are reactions we have to impressions of sense or other ideas. Since ideas are derived from impressions, and there are two kinds of impressions, ideas are derived from either impressions of sense or reflective impressions (or both).

Hume’s experimental method aims to explain the ideas we use in thought and language by tracing it to the impression from which it is derived. As he says, “’Tis impossible to reason justly, without understanding perfectly the idea concerning which we reason; and ’tis impossible perfectly to understand any idea, without tracing it up to its origin, and examining that primary impression, from which it arises.”

Thus, as we saw in chapter 1, Hume applies his method in moral philosophy to study the ideas of virtue and vice. He first asks what impression is common to instances where we judge that there is a virtue or a vice. He argues that since judgments of virtue and vice motivate, and reflective impressions are what influence the will to

202 T 1.3.2.4, 74-5; Don Garrett, Cognition and Commitment, 2.
action, virtue and vice are derived from the reflective impressions of moral approval or disapproval.

But Hume method does not end with such description. As he says in the introduction to the Treatise, his aim is to “endeavour to render all our principles as universal as possible, by tracing up our experiments to the utmost, and explaining all effects from the simplest and fewest causes.”

His method does not stop with finding the impression that gives rise to the idea, but explains the idea by pointing to the causes that bring about the impression from which the idea is derived. Thus when explaining the ideas of virtue and vice, Hume does not stop at pointing out the impressions of moral approval or disapproval that give rise to these ideas. Recall Hume’s explanation of virtue and vice: “An action, or sentiment, or character is virtuous or vicious; why? Because its view causes a pleasure or uneasiness of a particular kind. In giving a reason, therefore, for the pleasure or uneasiness, we sufficiently explain the vice or virtue.”

Hume concludes that if you wish to explain virtue and vice, you need discover the causes of the impressions that give rise to these ideas. Thus virtue and vice are explained by agreeableness or utility to self or others.

In the following I apply Hume’s experimental method to the idea of obligation, in three stages. In stage A, I bring examples of the phenomenon of obligation and draw a list of criteria for the kind of impression that unites these instances; in stage B I propose a motive that

---

203 T 0.8, xvii; see discussion in Hazony and Schliesser, “Newton and Hume”, 6-8.
204 T 3.1.2.3, 471.
205 “Thus, to take a general review of the present hypothesis: Every quality of the mind is denominated virtuous, which gives pleasure by the mere survey; as every quality, which produces pain, is call’d vicious. This pleasure and this pain may arise from four different sources. For we reap a pleasure from the view of a character, which is naturally fitted to be useful to others, or to the person himself, or which is agreeable to others, or to the person himself.” T 3.3.1.30, 590.
can be the reflective impression that we experience in these instances; and in stage C, I point to the causes of this impression which explain the idea of obligation. Applying this method offers an explanation of the idea of obligation in terms of the causes of the impression that gives rise to it.

2.2 Hume’s Method Stage A: Describing the Phenomena of Obligation

In order to apply Hume’s explanatory method to obligation we must first look at some instances of obligation. Different instances of obligation will share a certain impression, which led us to categorize them as instances of the same idea. In order to see this, we must first describe such instances. Let us first get clear on the phenomena, by giving three examples of individuals who act out of obligation:

➔ Sara and her brother Sam are waiting for the train home on a rainy Friday. Sundown approaches, the time when their mother will light candles to signal the beginning of the Sabbath. This is when, according to the Jewish tradition, you must desist from creative work and the use of technology for twenty-five hours. The siblings watch as the sky darkens, but no train arrives, and they realize that they will have to walk home, a three hour walk from downtown Manhattan to their home in the Bronx. Tradition prohibits taking the train on the Sabbath, and the walk ahead is not getting shorter as they feel bad for themselves. Sara urges that they should start their long walk since she is hungry and tired. But Sam’s legs hurt from his run last night, so he loiters on in conversation, and as the train finally approaches, he considers getting on. The train stops, the doors open, and Sam glimpses a friend from work beckoning to him excitedly. Obliged to keep the Sabbath, he tears himself away and pulls Sara in the direction of the stairs.

➔ Kate and Lily are best friends, and tell each other all their secrets. Kate is the only person in the world who knows how unhappy Lily has been this past year, since her husband began travelling for work. At first Lily would cry to Kate about her loneliness, and about fighting with her husband when he got back. But slowly her loneliness and bitterness lapsed into apathy. Lily spends a lot of time with Kate and her husband, and through them she meets their friend Jim. Kate had never heard Lily excited about another man, until Lily started talking about Jim. Jim is an easy-going guy, and who has been on his own for a long time. Lily’s husband would never know, if she became involved with Jim, since he is never
around. But a few days after Lily starts talking about Jim she stops coming over, and she eventually confides in Kate that she had decided to avoid Jim completely because of her obligation to her husband.

Jane is a successful lawyer, happily married with two children. Over the past few years, she has listened with helpless concern to reports from family and friends about her sister, Tess, who is falling into alcoholism. She follows as Tess loses her job, and sometimes has long conversations with her on the phone. One day Tess calls her and tells her that she has no money and is being kicked out of her apartment. Although Jane lives in a comfortable apartment that could easily give shelter to her sister, she does not offer that Tess move in. She thinks offering Tess a place to live will encourage her sister to continue drinking and will harm her own children. Although Jane thinks that this course of action has the best consequences overall, she still feels a constant regret at this neglect of her sister. She knows she has an obligation which she is not fulfilling.

In the first example above, Sam and Sara both have an obligation to keep the Sabbath. They approve of keeping the Sabbath, and if you asked them why they keep the Sabbath they would most likely say that it is good to do so, they approve of this tradition. But in this moment of conflict, when they are motivated to seek rest, comfort, and the company of friends, what motivates them to keep the Sabbath is an obligation to keep the Sabbath, not approval of keeping the Sabbath. Similarly, Lily surely approves of fidelity, and would say as much if she were asked. But when she is angry at her husband and attracted to another man, what motivates her to remain faithful is not her moral approval of faithfulness (alone) but an obligation to her husband. Finally, Jane morally disapproves of giving her sister shelter, because she thinks it will encourage her sister to continue drinking. And yet although she morally disapproves of this act, she continues to feel regret that she is not fulfilling an obligation to care for her sister.

From the examples above we can elicit some criteria for the impression that gives rise to the idea of obligation. First, the impression that gives rise to the idea of obligation is a
motive. The impression of obligation has a motivating influence – it motivates Sam and Sara to walk home, and it motivates Lily to avoid Jim. It also motivates Jane, although other motives outweigh it. Its motivating influence is felt by Jane as a constant urge to action which is not acted upon. On Hume’s categorization, if an impression has a motivating influence, then it must be a reflective impression such as a desire or passion.

Jane’s case emphasizes why obligation must have its source in a motive. One could argue that obligation arises from whatever is the strongest motive to act, or whatever turns out to be the most approved motive. Jane’s case shows that even when obligations are not fulfilled, they continue to influence us. This means that even when they are not the strongest motivation, and do not overcome other motivations, we have a distinct impression that gives rise to the idea of obligation. This impression could be felt by Sam or Lily, even if they did not act to fulfill their obligation. Sam would have the impression that he is obligated to keep the Sabbath, even if he chose to get on the train, and Lily would have an impression of an obligation, even if she chose to betray her husband. This strengthens the conclusion that obligation arises from a distinct motive.

Second, the impression of obligation is noticed in instances of struggle. It is in these instances of struggle that we have the impression that we have an obligation to act according to our moral approval, which we would not be aware of in cases where we were not conflicted about acting according to our moral approval. This doesn’t mean that we don’t have an impression of obligation in cases where there is no conflict, but rather that when obligation and moral approval go hand in hand the distinct motive of obligation is harder to detect. Thus I conclude that since the instances where we are most aware of our obligation are instances of
struggle, the motive of obligation is one that is strong enough to influence our will in face of strong opposing motives.

Third, the motive of obligation will be a motive distinct from moral approval or disapproval. Obligation refers to an act taken in face of opposing motives, and most often when we are not motivated (or not sufficiently motivated) to act according to our moral evaluations. When we look at examples of actions taken out of a sense of obligation, the influence of moral approval or disapproval appears weak, but there appears to be another motive that influences our will to act according to our moral judgments.

Fourth, this motive must be able to motivate some actions that we morally approve of, but it does not need to motivate all such actions. This is because we don’t need to assume that all moral actions are obligatory.

Thus I conclude that in order to offer a solution to Hume’s problem of obligation we need to point to

a) An reflective impression that motivates (desire, passion or sentiment)
b) while withstanding opposing motivations (such as self-interest)
c) that is distinguished from moral approval and disapproval
d) and which motivates a subset of actions approved by the moral sentiments

2.3 Hume’s Method Stage B: Identifying Loyalty as the Common Impression

Now that we considered different instances of acting out of obligation, and we elicited criteria for the impression that is common to them, let’s propose a reflective impression that is the common impression to these instances of obligation. In this section I argue that the reflective impression that gives rise to the idea of obligation is the motive of loyalty. Loyalty is a
motive to act on behalf of those who are part of one’s extended self. In this section I show that loyalty is (a) a motivation to action (b) which is distinguished from moral approval and disapproval and (c) motivates a subset of approved actions while (d) withstanding opposing motivations. My claim is that there are many instances where moral sentiments are too weak to motivate moral action, just as Hume suggested. In these instances loyalty is the motive that influences our will to act and thus give rise to the idea of obligation.

First, let us recap what loyalty is and how it motivates action. Recall that in the previous chapter I argued that a person has an extended self, which includes individuals and groups. The relation of these people to the self is the basis for a distinct motive to act their behalf. Thus I defined loyalty as acting on behalf of the needs and interests of those who are part of the agent’s extended self when motivated by the fact that they are part of the agent’s extended self. Can loyalty be the motive to action that influences the will in the examples above? Let us consider how loyalty motivates in these instances.

In the case of Sara and Sam, each may approve of keeping the Sabbath or may approve of the set of actions required by the Jewish tradition as a whole. But in addition, they have a certain relationship to the tradition, it is theirs, and thus it is part of their extended self. We see this because when people point out the beauty or morality of the tradition, they feel pride, and when the tradition is defamed or when it is immoral, they feel humiliation. They feel these emotions, which are emotions we feel towards our selves, because the tradition is part of their extended self. This also means that they have a certain relationship to the needs or
requirements of the tradition. Thus Sara and Sam can be motivated to keep the Sabbath not only by their approval of such behavior, but out of loyalty to the tradition that instructs them to keep the Sabbath. Sara is motivated to give up her own physical comfort and walk for three hours, out of loyalty to the instructions her tradition gives her. Sam is motivated to give up a good chat with a friend, and perhaps a good work connection, because if he does not a part of himself will be neglected. His tradition, which is already a part of him, would be pulled away from him by this action. As his tradition becomes a little less his, when he refuses to fulfill its needs or act upon its instruction, his own sense of ability to bring together the parts of who he is and to succeed in strengthening them is weakened. He acts in a way that harms his own ability to extend himself in a successful way.

Lily’s husband is part of her extended self: she feels pride at his success and is humiliated by his failings, because the idea of him is part of the idea her self. Lily morally disapproves of betrayal, which not only harms another person but is an abnegation of a commitment. But this moral disapproval does not always strongly motivate her, when she feels anger towards her husband and desperation to find someone who will appreciate her and show concern for her. In such instances, there is another motive that motivates Lily to remain faithful. She is motivated to care for her husband and fulfill his needs not only out of moral approval for such an act, but because he is part of her extended self. When she succeeds in assisting him, she feels that her extended self is succeeding, and that she is a capable person who can bring about the improvement of her own extended self. If she betrays him, she will be harming a person who

---

206 It may not be clear here how a tradition can have needs. My view is that any particular object can be strengthened or weakened, and thus that it has needs. This will be discussed in greater detail in section 2.6 below and in chapter 4, where I argue that we can have loyalty to a group, which is a particular entity, but we cannot have loyalty to a rule, which is an abstract entity.
is part of her extended self. Her own sense of self will crumble, because a part of her will be torn from her. Unlike her husband’s death from cancer, such a loss will be one she causes. Thus, since she will be causing this crumbling of her extended self, it will be an expression of her lack of ability to strengthen her extended self.

2.4 Loyalty Motivates a Subset of Moral Actions

Above I claimed that one of four criteria for the impression that gives rise to the idea of obligation is that it motivate a subset of morally approved actions. Clearly the examples I chose are meant to focus our attention on examples where loyalty motivates actions we approve of. But does loyalty indeed motivate a substantial subset of morally approved actions? I here show that loyalty does in fact motivate a large subset of moral actions on Hume’s moral theory.

Although from philosophical discussions of trolleys and bridges it may sometimes appear that many of our moral actions are directed towards people we never met, this is not the case in our lives. Many, if not most, of our moral actions take place in our ongoing relationships. The more we come in contact with someone, the more we will have opportunity to act morally in our relationship with that person. This fact has given rise to renewed interest in associative duties. But these discussions of the normative status of relationships usually do not characterize associative duties as more central to morality than others. Associative duties are presented as merely another moral reason for action.207

Hume’s theory gives a more prominent place to ongoing relationships in the judgment of a person’s character, and derivatively, of approved actions. As Hume scholars have shown, Hume thinks that we make our moral judgment from a common point of view, from which we approve or disapprove of people’s characters in a non-self-interested way. But interestingly, Hume thinks that we judge other people’s characters by considering their relationships with their close circle of associates. We judge whether the person is motivated to act with courage, wisdom, beneficence and fairness towards those he or she has an ongoing relationship with, what has come to be called Hume’s “close circle”. Dorothy Coleman and Erin Frykholm have argued that for a Hume, we primarily judge a person’s character based on how this person is motivated to treat those people who he or she comes into regular contact with, such as family, friends or co-workers. On this view, if a person acts beneficently in a passing interaction with a person she will never see again, this will not influence the judgment of her character. This is not so much because Hume views helping strangers as irrelevant to morality, but because, as we saw in chapter 1, Hume thinks morality is primarily concerned with a person’s character and character is judged by ongoing rather that individual instances.

My claim is that if we primarily judge other people based on how they are motivated to act in their ongoing relationships, then this consideration is also central to our moral judgment of what we ought to do. Just as our approval of other people is primarily a judgment of how they

---

208 Dorothy Coleman, “Partiality in Hume’s Moral Theory”, Journal of Value Inquiry, 26 (1), 1992, 95-104. Erin Frykholm points out that the extent of the close circle will change, depending on the person’s circle of influence. Thus the close circle of a politician may include their constituency, many more people than the close circle of a non-politician, since the politician influences their lives in dramatic ways. See Erin Frykholm, “Associative Virtues and Hume’s Narrow Circle”, Pacific Philosophical Quarterly, 97 (4), 2016, 612-637. See also Geoffrey Sayre-McCord, “On Why Hume's ‘General Point of View’ Isn't Ideal–and Shouldn't Be”, Social Philosophy and Policy 11 (1), 1994, 202-228.
are motivated to act in their relationships with those in their close circle, so our moral judgment of what we ought to do will be heavily influenced by the needs and interests of those in our close circle. We may not be aware that we make the calculation, but whether someone is part of our close circle or not is an important consideration in our moral judgment of what we ought to do. This does not mean that we do not approve or disapprove of actions when they are directed towards strangers, but rather that these are not as central to how we judge the character of others. With regard to our judgment of obligation, this is not where we have the original impression that gives rise to the idea of obligation.

The challenge that arises at this point is that we do also have obligations to people who are not in our close circle. Can loyalty motivate us to keep our promise to a stranger, or to refrain from stealing from a stranger? I will argue that impersonal obligations come from the extension of our loyalty to a group, and thus are modeled on the obligations of loyalty to individuals in the close circle. How this happens psychologically will be discussed in chapter 4, when we consider how obligations arise from Hume’s the artificial virtues. For now it is sufficient that loyalty does motivate a subset of morally approved actions – at least those that pertain to acting on behalf of those individuals who are in our close circle. What this amounts to is not that our obligations are limited to close relations and associates, but that these obligations are the models for what moral obligation is. This is because loyalty is the motive that gives rise to the idea of obligation, and loyalty to those in the close circle is our model for loyalty.
2.5 The Strength of Loyalty

So far we’ve seen that loyalty is a distinct motive, which motivate separately from moral approval and disapproval, but also motivates a subset of morally approved actions. In this section I show that loyalty is a strong motive to action, by revisiting some of the arguments made in chapter 2, in my characterization of Humean loyalty.

We started out with Hume’s judgment internalist view that moral sentiments explain the influence of moral judgment on the will to act. I objected to this claim, arguing that moral sentiments influence us in too many instances that do not require action, and influence too weakly to influence us to act. The above instances of obligation exemplify why moral sentiments on their own cannot explain the influence of moral judgment on the will, because they are examples in which moral approval comes in conflict with other strong motives. In the case of Sara and Sam, their approval of keeping the Sabbath comes in conflict with desire for rest and the love of company. In Lily’s case, the struggle is much more dramatic, since her moral approval comes up against strong motives of hatred towards her husband and desire for a new romantic interest. In Jane’s case, moral approval of caring for her sister comes up against loyalty to her children. In all of these cases, I suggest, we see examples of the claim Hume made, that moral approval and disapproval are weak motivations to action. I argued that in these instances loyalty is the strong motivation to act in accordance with moral approval.

In order to understand why moral sentiments are weak motivations to action, while loyalty is a strong motive to action, we need to revisit the argument I made when I distinguished Humean loyalty from Humean sympathy. There I showed that sympathy only brings about a weak motivation to action, because it associates a sentiment with the self. This sentiment, like
many human sentiments, is only motivating so long as it is present, and comes and goes like other sentiments. Loyalty, on the other hand, is generated by associating another person with the self. This association, because it is an association between ideas, requires a more constant association, since objects do not change all the time. It is because the association between objects that make up the self is more constant than the appearance of moral approval or disapproval that the motive of loyalty is more constant. This constancy gives the motive of loyalty more influence on the will, and thus makes it a stronger motive.

2.6. Hume’s Method Stage C: Causes of Loyalty and Explanation of Obligation

Up until now we’ve been looking at the impression common to instances of obligation, which is the source of the idea of obligation. As I mentioned in the section on Hume’s methodology, Hume’s view is that advance in science is made by explaining ideas. An idea is explained by looking at the instances of an impression, and pointing to its causes. As we saw in chapter 1, Hume applies this method of explanation when he argues that our idea of virtue comes from our impression of moral approval, and then points to the properties of character traits that cause this impression of approval – the utility or agreeableness to self or others. In this section I offer an explanation of obligation by pointing to the causes of the impression of loyalty, which is the impression that gives rise to the idea of obligation. Just as virtue is explained by agreeableness and utility to self and others, obligation is explained by the needs of the extended self.
First, let us understand why Hume thinks this method offers an explanation. This search for the causes of the impression allows us to understand the idea, because it explains what its causes are and therefore what context it arises in. What is important to point out is that the idea has its source in an impression, not in the causes. Thus, for example, we cannot understand the meaning of virtue by grouping together character traits that are agreeable or useful. It is only once we have an idea of virtue that we can go back to the instances that gave rise to the idea and figure out what are the characteristics of these instances that cause the same impression. The idea has its source in an impression, but the impression on its own doesn’t explain the idea.

Once we explain an idea, by pointing to its causes, we can also check our use of this idea. If we find the causes of the moral impression, then we can check if all the instances in which we use the idea of virtue are appropriate. Thus pointing to the causes of virtue and vice allows us to critique people’s use of the ideas virtue and vice, thus critiquing their list of virtues and vices. This is because once we’ve found the causes of moral sentiments, we can argue that some instances of moral sentiment are inappropriate if they are not produced by the causes of these sentiments. This is what I take Hume to be doing when he argues that the monkish

---

209 For more detailed discussion see Hazony and Schliesser, “Newton and Hume”.
210 Garrett, Cognition and Commitment, 198-199.
211 The view that there could be wrong instances of sentiment, because they are the wrong response to their causes, can seem strange to us. This is because we think of emotions as completely personal, and thus as not having appropriate and inappropriate causes. But this is not how emotions are viewed in research of emotions, where psychologists and philosophers study the causes and effects of emotions. Moreover, we can think of daily examples where we clearly think there are appropriate and inappropriate instances of emotions. If a child becomes angry and throws all his food on the floor because we put a drop of ketchup on his plate, we think that this is an inappropriate reaction. This is not because we think that the child can never become angry in instances where we would not be angry. Anger is defensive, and erupts when we feel that we are in danger. Thus we would find it appropriate if a child gets angry when another child hits him, even though we would not respond in the same way ourselves. But we still think that the child’s anger, in the ketchup case, is inappropriate, because we think that there is no danger to the child at all.
virtues are no virtues at all. In a similar way, if we can point to the causes of loyalty – which I have argued is the impression that gives rise to the idea of obligation – then we can have an explanation of the idea of obligation. By pointing to the causes of loyalty, we can point out what can or cannot be called an obligation.

What are the causes of loyalty then? The causes of loyalty are the needs of those who are part of the agent’s self which can be filled, at least partially, by the agent. Let us go back to the example of Lily avoiding new romantic relationships out of loyalty to her husband. In this example, the impression that gives rise to her idea of obligation is the motive of loyalty she has to her husband, which motivates her to overcome her anger and romantic interest in another man. But what is the cause, which brings about this motive of loyalty, in the particular situation? There is something that causes the motive of loyalty to appear. The cause of Lily’s motive of loyalty, in this particular instance, is her husband’s need of protection from the harm she could cause him. Her husband’s pending devastation and humiliation cause Lily to be motivated by loyalty to avoid such a situation. All instances of loyalty have similar causes, the needs of those who are part of the extended self.

Hume writes “Celibacy, fasting, penance, mortification, self-denial, humility, silence, solitude, and the whole train of monkish virtues… We observe, on the contrary, that they cross all these desirable ends; stupify the understanding and harden the heart, obscure the fancy and sour the temper. We justly, therefore, transfer them to the opposite column, and place them in the catalogue of vices; nor has any superstition force sufficient among men of the world, to pervert entirely these natural sentiments” (M 9.3, 270). Hume only says that all people would agree that the monkish virtues are no virtues, but he doesn’t address the fact that some people do think that they are virtues. One can argue that Hume didn’t see the agreeability or utility of these character traits. But this is beside the point. Hume’s point is that his method allows us to understand virtue by tracing it to its original impression, pointing out the causes of this impression, and then checking our own impressions to make sure they are caused by the appropriate causes. Hume’s claim is that if virtue is explained by the characteristics of being useful or agreeable to self or others, and the monkish virtues do not have these characteristics, then they cannot be virtues.

Does this mean that a person who is motivated by loyalty must fulfill all the needs of any person who is part of their extended self? This seems to again give rise to too many obligations, and not to reflect the instances in which we are motivated by loyalty. Allen points out that not all relationships require full and constant filling of all needs, but rather there are specific needs that are included in the kind of relationship it is. R. T. Allen, “When Loyalty No Harm Meant”, Review of Metaphysics, 43, 1989, 283.
Since the causes of loyalty are the needs of certain objects, the objects that we can be loyal to cannot be objects that have no needs we can fulfill. We can only be loyal to things whose situation we can improve in some way:\textsuperscript{214} we cannot be loyal to a planet in outer space or to a volcano, who have no needs we can fulfill. We also cannot be loyal to an abstract principle, which has no needs. Only particular objects with needs can be betrayed, when we do not act on their behalf. Only failing to fulfill such needs can explain the harm to the agent when they are disloyal – the tearing apart of the extended self when a person fails to act upon the need of someone or something that is part of their extended self. This also means we can only be loyal to particular objects.\textsuperscript{215} Only a particular object can be part of the extended self. We cannot be loyal to a principle, since an abstract principle cannot be part of a particular self. Finally, loyalty is always acting on behalf of someone or something that we already have some kind of relationship to. Although we can be kind, gracious or vicious to someone we have no relationship to, we cannot be loyal to someone we have no previous connection to.

If the causes of loyalty are the needs and weaknesses of those who are part of the extended self, and loyalty is the impression that gives rise to the idea of obligation, then obligations can only be to those who are part of the extended self. This description of the causes of loyalty gives us the meaning of obligation: obligation is an action that fulfills a need of something or someone who is part of our extended self. As such, an obligation is (1) an action that fulfills a need (2) of a particular object that is (3) already part of the extended self of the

\textsuperscript{214} Allen, “When Loyalty No Harm Meant”, 293.
\textsuperscript{215} Ibid, 287: “the object of loyalty must be either a concrete (and personal) entity to which one stands in some sort of relation, acknowledged at least on the one side, or a pledge of service to something which is not a concrete entity, such that there is, in either case, a promise or expectation of some sort of service, which can therefore be given or refused. Loyalty, then, is loyalty to something already in existence or performed – something there.”
agent. This means that if we say that someone has an obligation to do something for someone who is not part of their extended self, we are misapplying the idea of obligation.

This part utilized Hume’s experimental method to discover the psychological source of obligation. It argued that the idea of obligation has its source in the impression of the motive of loyalty, and we can explain what obligation is by discovering the causes of loyalty. In the following section I unpack the implications of this claim for a Humean internalist theory of moral judgment.

3. Rethinking Humean Moral Judgment

What I proposed to show in the previous sections was that moral approval, as Hume describes it, cannot have the kind of motivating influence that we find in instances of moral obligation. Thus I proposed that we must distinguish between moral approval and moral obligation. While Hume’s moral sentiments can explain our moral approval and disapproval, they cannot explain our moral obligation. I argued that loyalty is the impression that gives rise to the idea of obligation, because it is the motivation that comes together with moral approval in instances where we act out of moral obligation. Since the causes of loyalty are the needs of those who are part of our extended self, obligations are actions that fulfill the needs of those who are part of our extended self.

This suggestion raises a concern if it is to be understood as a Humean position. Loyalty is distinct from, and therefore external to, moral judgment as Hume presents it. This is because Hume claims the source of moral judgment is in the sentiments of moral approval and
disapproval, while loyalty is a different motive. Rather than arguing that a Humean must give up on judgment internalism, in this part I offer a new version of Hume’s judgment internalism. I suggest that we can distinguish between the judgment that a character trait is virtuous and the judgment that an action is obligatory. While we only need one impression – the impression of moral approval – to give an internalist account of our judgment that a motive or action is virtuous, we need two impressions – the impression of moral approval and the impression of loyalty – to give an internalist account of our judgment that an action is morally obligatory. Thus I offer an internalist account of the moral judgment of obligation which is made up of a calibration of moral approval and loyalty.

In the following I argue that we calibrate moral approval and loyalty in judging moral obligation because moral obligation is the calibration of impartial good with the partiality needed for human action. In the first section of this part I argue that we should view obligation as a unified idea arising from a single impression. This will allow me to treat obligation as one axis of our moral judgment. In the second section I argue that loyalty is not merely one of many considerations for obligation, but rather is central to our judgment of obligation, because obligation to act always has its source in a partiality one has towards the situation. Finally, in the third section I describe how we calibrate the impartiality of moral approval and the partiality of loyalty in judging that we have a moral obligation.

3.1 Unity and Pluralism in the Humean Idea of Obligation

My argument earlier in this chapter aimed to show that for empirical reasons, a Humean should accept the Rossian view that the right is not reduced to the good and that these
are distinct concepts. In what follows I argue that on a Humean view, although we have obligations in different circumstances, there is one impression and one cause that together explain what obligation is. Unifying the idea obligation in this way will enable me to then explain how obligation can be calibrated with moral approval in our judgment of moral obligation.

While I’ve tried to show that a Humean should accept Ross’s distinction between what is right and what is good for empirical reasons, a Humean will not share Ross’s view of the content of each of these concepts. According to Hume, the idea of moral goodness arises from the approval we feel about motives and systems of conduct from a common and general point of view, and it is explained by the utility and agreeableness of motives and systems of conduct to self and to others. Similarly, a Humean holds that the content of the idea of obligation, and therefore the nature of obligation, will be explained by pointing to an impression and its causes.

As I argued in the first chapter, Hume holds the view that the impressions of moral approval and disapproval are caused by both motives and systems of conduct, and therefore that moral judgment is pluralistic in the sense that moral approval comes about in different circumstances and in response to different objects. But this does not change the fact that the idea of virtue, for Hume, is traced to a particular impression of moral approval that we have towards both motives and systems of conduct that are useful or agreeable. It is the usefulness and agreeableness, and not the differing circumstances, which explain the judgment of virtue.

---

216 Ross explicitly distinguishes his view of the good from a view that characterizes goodness by approval. *Right and the Good*, 75.
The same, I contend, is be true for the impression that gives rise to the idea of obligation. There is a single motive, which appears in the different instances of obligation, and this is the impression of loyalty. If we look at the characteristics external to the agent, there is no commonality between the circumstances that give rise to obligations in Sarah’s and Lily’s cases. There is little similar in the circumstances of Sara who is obligated to keep the Sabbath and Lily who is obligated to refrain from engaging in a new romantic relationship. But if we look at the reflective impressions that Sarah and Lily both have, they are both motivated by the motive of loyalty, and thus there is a common impression uniting these different circumstances.

As we saw, for Hume an explanation of an idea is achieved by pointing out the causes of the common impression. Although there are different circumstances in instances of obligation, the impression of loyalty has certain causes. Thus we can explain obligation in terms of the causes of loyalty. Loyalty is caused by the needs and interests of those who are part of the agent’s extended self. The motive of loyalty is caused by different causes, in Sarah’s and Lily’s cases: Sarah loyalty is caused by the possible denigration of her tradition, Lily’s loyalty is caused by the possible suffering of her husband. But both of these causes are of the kind of things that cause loyalty: the needs and interests of particular objects that are associated with the self.

Notably, this explanation of obligation can lead to conflicting obligations in many situations. Obligation is not what one must do, all things considered, on this Humean account. Obligations are actions that fulfill the needs of those who are part of the extended self, and these can be numerous and conflicting. This is in line with Hume’s pluralism regarding virtue. Just as Hume thinks the demands of virtue can conflict, on his view demands of obligation can also conflict. I see no source for distinguishing between obligation and *prima facie* obligation, on
Hume’s view. Although I critique the close connection Hume draws between virtue and obligation, I retain this pluralism on my view of obligation.

3.2 The Function of Obligation

In the previous section I argued that for Hume there is one impression that gives rise to the idea of obligation. In this section I explain the function of this impression, meaning what role it plays in our moral judgment. This is not a pluralist view, like the now-common pluralist view that hold that we have both agent-neutral and agent-relative reasons for moral judgment. Instead, I argue that there is a certain relationship between these two considerations. I argue that moral approval and disapproval are prioritized by the limiting influence of loyalty, meaning by the extension of the self of the agent in question, and thus that both components are necessary for the moral judgment of obligation.

Many theories of moral judgment focus on the agent-neutral aspect of morality, and therefore view moral judgment itself as the application of agent-neutral rules or the weighing of agent-neutral reasons. There is no stronger example of this view than Tom Nagel’s *The View from Nowhere*, in which he argues that moral judgment is characterized by its agent-neutral position. As Nagel puts it, “Values are judgements from a standpoint external to ourselves…They tell me how I should live because they tell me how anyone should live.”

---

217 Tom Nagel, *The View from Nowhere*, Oxford University Press, 1986, 135. Nagel argues that agent-relative reasons can be viewed from the agent-neutral perspective in chapter VIII. For a critique of the view that the moral point of view is always impartial see John Cottingham, “Partiality, Favouritism and Morality”, *Philosophical Quarterly*, 36, 1986, 357-373.
In the last few decades some normative ethicists have argued that there is something missing from this picture of moral judgment. David Brink has pointed out that “most of us think that morality also includes demands of partiality owed to associates,” and that such agent-relative or relationship-relative obligations are not reducible to agent-relative ones. A similarly pluralist view can be found in Andrew Oldenquist’s suggestion that loyalty is one of three sources of obligation, alongside self-interest and agent-neutral moral approval and disapproval. I agree with the motivation of such theorists to include agent-relative considerations as components of our moral judgment. The problem with these pluralist theories, that place agent-relative reasons alongside agent-neutral ones, is that they fail to explain the relationship between the neutral and the relative aspects of moral judgment. Brink considers a few options, but is left with the pluralist view that both have normative force, and therefore a role in our moral judgment, without explaining any clear priority or relationship between them. On such pluralist views, we cannot offer a unified theory of moral judgment as characterized by an agent-neutral point of view, as suggested by Nagel, but we also don’t have an alternative characterization of moral judgment.

The difficulty of characterizing moral judgment on these pluralist views leads me to reject the position which characterizes agent-relative and agent-neutral considerations as different reasons for moral judgment. Instead, I argue that both agent-neutral considerations and agent-relative considerations are relevant in every moral judgment, but that these considerations play a different function and therefore interact in a particular way. Every moral judgment of

---

219 Ibid, 162-163.
221 Brink, “Impartiality and Associative Duties”, 169-172.
obligation has two components: the agent-relative component of loyalty, which prioritizes those who are part of the agent’s self over others, and the agent-neutral component of moral approval, which evaluates the virtue or vice of motives and the actions that come from them.

My claim is that when we make a moral judgment, we bring our agent-neutral moral approval and disapproval to bear on our loyalties, which are agent-relative. The agent-relative considerations that come from our loyalties prioritize among on our agent-neutral moral approvals, directing us to act on those most relevant to us, and it is this function of prioritizing which make obligation a strong motivation for action. On this view, all obligations come from loyalty, because there is only one psychological source that gives rise to the idea of obligation, and this is the impression of the motive of loyalty. The motivating influence of obligation comes from its agent-relative character, and not from its agent-neutral nature.

I want to distinguish this suggestion from the monist option Brink mentions, the option that says that only personal associations give rise to obligations.222 The problem with this view is that it leaves agent-neutrality out of the picture of moral judgment completely.223 I retain Hume’s view, that moral approval and disapproval are agent-neutral evaluations. As Hume says, we do in fact morally approve of the character of people we have never met and even the character of our enemies. My emendation of Hume’s view is that I hold that we have two components of moral judgment: moral approval and loyalty. While one component of moral judgment – moral approval and disapproval – is agent-neutral, there is another component that is

222 Ibid, 169.
223 Brink’s other critique, that we do have impartial obligations, will be addressed in chapter 4.
agent-relative. While our moral approval and disapproval apply to all humans everywhere,\(^{224}\) our judgment of moral obligation distinguishes human beings based on their association to the agent.

The agent-relative nature of obligation is missing from Hume’s account of obligation, just as it is missing from many accounts of obligation. But since Hume claims to be offering a psychological source for obligation, we can faithfully rethink his suggestion based on his own premises. Hume is looking for an internalist account of moral judgment, an account that can explain how moral judgment is motivating. As I argued previously, what is missing from Hume’s picture is the influence that prioritizes among our moral evaluations. When a person acts morally they view the moral judgment in relation to their other motivations for action. We need to be able to explain why a particular moral judgment will motivate, given the particular agent’s other motivations for action. We need to be able to distinguish approving of certain actions from judging that they are obligatory.\(^{225}\) My claim is that the moral judgment regarding an obligation is not agent-neutral, because it is a judgment about the priority a particular instance of moral approval or disapproval has for the agent. This priority is not a claim about the preferences of the agent, but it is a claim about the priority this moral approval or disapproval has for motivating the agent to action.\(^{226}\) This priority setting must take into account the motivational scheme of the agent, and thus is agent-relative.

\(^{224}\) Though it may be applied with some sensitivity to age, knowledge, or societal norms.

\(^{225}\) This relational aspect should not be confused with egoism. The fact that obligation is relational doesn’t mean that obligation arises from the agent asking “what is in this for me?” Hume clearly thinks that motivations can be non-egoistic, and argues that moral judgment is not a judgment regarding what is good for me.

\(^{226}\) This relational connection should not be confused with subjectivism. It is not up to the agent to decide what his or her obligations are, this would be subjectivism. But the existence of obligation is tied to psychological make-up of the individual, and therefore is agent-relative.
What we do, when we judge that an action or set of actions is obligatory, is judge how much priority it has for us. This judgment is not a judgment of how much we want to perform a certain act, for this would just be the question of how much we are currently motivated to act in this way. The question is how important this action is, given the associations that make up our self. The priority of the moral judgment of an action is not decided by how great our approval of it is, nor by how much we want to act in this particular way, but rather by how closely the recipients of the actions are associated with the self. When we judge ourselves to have a moral obligation, we do not only judge whether the outcome will be good, but also whether it will be good for those who are part of our extended self. The fundamental question in obligation is who am I loyal to, and how will they be effected by my actions. It is this consideration, together with our moral sentiments, which defines our moral obligations.

We can now see how on my view we can distinguish between the role of moral evaluation that comes from the moral sentiments, and the role of judgment of moral obligation that combines the moral sentiments with loyalty. As Hume argued, moral approval and disapproval allow us to evaluate morally in a much broader spectrum of cases than the ones that we act upon. Although moral approval does motivate, it motivates with regard to many situations that are beyond the scope of the agent’s actions, and therefore it only motivates weakly. This is not useless for the agent, because it opens for consideration many situations which the agent is not facing but may come to face. Thus moral approval weakly motivates us to assist those in need in other times and places, and moral disapproval weakly motivates us to fight those who are cruel in a novel we are reading. Moral approval and disapproval thus allow us to discuss matters of moral concern with others, give and receive advice, and consider potential cooperation. But in
many of these instances, the weak motivation of the moral sentiments does not motivate strongly enough to produce action, and it does not need to.

Moral obligation, on the other hand, dictates how the agent should act. In order to fulfill this role, it must carve out a limited number of moral judgments that the agent will act on. As such, obligation raises an additional question to that of moral approval, namely: what is the relation of this moral approval or disapproval to the agent in question? This question can be asked in a first-person “what should I do?” or in the third person “what should she do?”, but in both cases it is a relational question: what is the relation between this moral approval or disapproval and the agent in question? While moral approval and disapproval evaluate all people, on Hume’s view, since they are felt from the common point of view, the judgment of an obligation is relative to an agent because it assesses how this moral approval or disapproval should motivate a particular person to action.

My suggestion is that we combine loyalty and the moral sentiments by carving out a subsection of morally approved actions that are actions on behalf of those who are part of one’s extended self. These actions we view not only as morally approvable, but as obligatory. The

---

227 Ross thinks duties are always to act in a certain way, not to have a certain motive. Ross, Right and the Good, 22-23.

228 Ross acknowledge the personal nature of obligation. “The essential defect of the ‘ideal utilitarian’ theory is that it ignores, or at least does not do full justice to, the highly personal character of duty. If the only duty is to produce the maximum good, the question who is to have the good… should make no difference to my having a duty to produce that good. But we are all in fact sure that it makes a vast difference,” Right and the Good, 22. Ross emphasizes that who receives the good matters to who needs to achieve it. But I add that if what matters to my duty is who will receive the good, then what stands in question is what is the relation between the receiver of the good and myself, the agent. See Ross, ibid, 38: “To make a promise is not merely to adapt an ingenious device for promoting the general well-being; it is to put oneself in a new relation to one person in particular, a relation which creates a specifically new prima facie duty to him, not reducible to the duty of promoting the general well-being of society.”

229 See Brink’s reference to Hume in his characterization of the moral point of view as impartial, Brink, “Impartiality and Associative Duties,” 154.
reason why we view them as such is that we think that a person in such a relation to the action will not only be motivated by moral sentiments but also by loyalty.

I am not suggesting that loyalty does in fact motivate in every instance where someone is obligated. A person is not always motivated by loyalty to their close associates. What I am suggesting is that we define obligations as instances where, given that someone’s self is extended to include others, a person could be motivated by loyalty. Thus when we say that someone ought to feed their children, or someone ought to visit a sick friend, what we are saying is that given that children or friends are associated with that person’s self, there is a motive of loyalty that could join moral approval in motivating this person to such actions. What loyalty adds here is that this motive is present for the individual under discussion because children or friends are part of this agent’s extended self. Thus the motive of loyalty is available to them.

My claim is that rather than viewing the identity of people as external to our moral judgment, we actually view it as one of the two factors that together form our judgment that we have, or someone else has, a moral obligation. This is because moral judgment of obligation is not merely a judgment that good could be done by someone acting in a particular manner in a particular situation. It is a judgment that someone in particular must act to achieve this good. Thus it must be a judgment of how this good is related to the person in question. Obviously many actions that are moral are not ones in the self-interest of the agent obligated. But my claim is that the actions that are obligatory are always on behalf of those who are part of the extended self of the agent in question.

To conclude, on my view agent-relative reasons for action, such as those that arise from relationships, are not merely one kind among many kinds of considerations that we are
susceptible to when we make a moral judgment. Rather loyalty in our relationships give us the framework to which we apply our moral approval and disapproval. Loyalty and moral sentiments are both factors of the moral judgment that an action is obligatory. Loyalty, the single impression that gives rise to the idea of obligation, is most easily distinguished in personal relationships, where the outcomes for the recipients are more easily detected. Thus personal relationships are the model which we use, to understand what our moral obligations are. As we shall see in chapter 4, this model can be extended to our relationships with groups of people, and thus give rise to impersonal obligations.

3.3 Calibrating the Two Axes of Moral Judgment

In the previous section I argued that moral judgment is made up of two components, and that the function of moral obligation is to prioritize which of our moral evaluations to act upon. But this description did not make it clear how the two components of moral judgment interact. In this section I offer a sketch of the way in which we calibrate the two components of moral judgment. I suggest that the two components can be placed on two axes, which allows us to calibrate the moral approval we have for a certain motive or system of actions with the extended self of the agent moved to act.

Ross considers the question how we weigh different *prima facie* obligations, but he doesn’t offer us any calculus of the kind we find in utilitarian theories. All he says is that over time, we find that we feel that some obligations are “more urgent” than others.\(^{230}\) “We come in the long run, after consideration, to think one duty more pressing than the other, but we do not

\(^{230}\) Ross, *Right and the Good*, 23.
feel certain that it is so.” Ross thinks that this indeterminacy does not make his theory any less strong than that of utilitarianism, which claims that some goods are just better than others. In what follows I build on Ross’s claim that there is no equation by which we calculate our moral obligation, but that there is a process of weighing that goes into deciding how we ought to act. On my suggestion, this process of weighing does not all occur in the moment of action. This is because one axis of moral judgment, loyalty, is not decided in the moment but is rather defined over time.

3.3.1 The Two Axes

In the previous section I argued that there are two components of moral judgment. I now suggest that these components can be placed on two axes. Moral approval and disapproval can be placed on an axis because we can approve or disapprove of motives and actions to a lesser or greater degree. So if we approve of the motive of beneficence, we evaluate actions of beneficence as virtuous, we evaluate giving five shekels in charity as less virtuous than staying awake all night to console a sick child. I do not think we morally approve of some virtues more than others. But we do evaluate actions that are derived from different virtues as more or less virtuous. So I do think we judge that giving five shekels in charity is less virtuous than the courageous act of calling out a colleague who is belittling another colleague. Regardless of this question, at the very least we are able to place actions that are approved of because they stem from the same virtue on an axis from smaller to greater.

231 Ibid, 30-31. This judgment is like judging beauty “In this respect the judgment as to the rightness of a particular act is just like the judgment as to the beauty of a particular natural object or work of art… Both in this and in the moral case we have more or less probable opinions which are not logically justified conclusions from the general principles that are recognized as self-evident,” Ibid, 31.
The other component of moral judgment, loyalty, can also be placed on an axis. Loyalty stems from the association of a beneficiary with the idea of the self. Here, too, we can put the actions on an axis that indicates how closely the beneficiaries of the action are associated with the self. On this axis the very same action, say staying up all night with a sick child, will be weighed more heavily when it is performed on behalf of one’s own child, and will weigh a little less when performed on behalf of one’s niece or a friend’s child, and will weigh much less when performed on behalf of a stranger. But notice that configuration of actions on this axis will depend on the self of the agent. So for a woman who takes on a babysitting job, or is a nurse in a hospital, caring for a child she doesn’t know will no longer be less of an obligation than caring for a friend’s child, and may even come close to weighing as much as the obligation to her own child. This is because her job relates the stranger to her self, and thus causes her to act on a motive of loyalty and obligation.

The calibration of moral approval with loyalty frees us from acting on whatever the greatest virtuous act is, no matter the long-term effects on the self. Loyalty, as we saw, is motivated by the needs of those associated with the self. But the self is only extended as far as it can act without disintegrating. We do not act on behalf of different people every day, because then there would be no association created between the beneficiaries and the self, and the self would not extend so that it could benefit from gaining in pride and sense of accomplishment.

---

232 Recall that distinguishing among beneficiaries is distinctive of obligation, according to Ross. See Ibid, 22.
233 A nurse is not obligated by a relationship to each of her patients. This view of a job as giving rise to obligations is not clear on the account of obligations are arising from personal relationships that I discussed in this chapter. Obligations arise from her job as a nurse, which fits better with the psychological explanation of impersonal obligations I offer in the next chapter. The job extends the nurse’s self to include the group of those who are sick, and loyalty to this group underlies her obligation to particular sick people. Alternatively, her job extends the nurse’s self to include the medical staff she is a part of, and this loyalty underlies her obligation to particular sick people.
from these actions. This is because acting on behalf of others also has costs for the narrow self, and thus must be returned in another way. Acting on behalf of others with no returns for the self would lead to the exhaustion of the agent. Loyalty ensures that we are most strongly motivated to act for those whose success and improvement also strengthens our own sense of self accomplishment.

3.3.2 Calibration in Judgment of Action

Now that we’ve seen that both components of moral obligation can form independent axes, we can discuss how these axes are calibrated. My suggestion is that rather than morally judging what action to take based on maximizing one of these components, we calibrate both. So it is not the case that we are morally obligated to act in the way that can be shown to be the most virtuous, nor is it the case that we are morally obligated to act in the way that is most loyal. Rather when we are conflicted between different actions, we consider them both for their virtue and for their loyalty components. This judgments points out our obligations. We are morally obligated by actions that are the strongest on both of these components together. Neither of the components can be completely transcended, so we cannot have a moral obligation that is evil, nor can we have a moral obligation that is not an action of loyalty to someone. Let’s consider an example to illustrate this point.

I am a nurse in the hospital and I get a call that my son is sick at home. I need to choose between staying in the hospital to care for my patients, and going home to care for my son. If my son has a cough, and my husband is home to take care of him, then I judge that my moral obligation is to care for my patients, since I have a loyalty to them and their need is greater
than that of my son. But if my son is very sick, and my husband is not home to care for him, then I judge that the close association of my son with my self renders it a greater obligation than my moral obligation to my patients. What is important to notice in this example is that there are two considerations: the need of the sick, which measures how beneficent or virtuous the actions is, and the closeness of the sick to myself, which measures the urgency to me of this action. It is less important to me whether you agree with me that my son should weigh more than my patients on my loyalty axis, or whether you agree with me that my son needs me more than my patients if his father is not there. What is important is that there are two axes – the axis of virtue and the axis of loyalty, and that both are components of judging obligation.

This account of moral judgment as having two components entails a particular normative view. It entails that we do not have a moral obligation to complete strangers, since we don’t have any preexisting association that can give rise to the motive of loyalty. We also do not have moral obligations to act in ways that are morally reprehensible, in order to help those who are part of our extended self. But most of our moral deliberation is in neither of these extremes. This is because most of the time, when we are deciding what to do, we are torn between the needs of parts of our extended self. Our extended self, which includes associations to individuals, groups, and society, leads to conflicting needs and interests. Most of the time, when we are deciding what to do, we are torn among our loyalties. Much of our moral deliberation is not about which actions we approve or disapprove of, but how to harmonize among competing loyalties. We do this by finding common ground among those we are loyal to, just as we try to find common ground between our narrow self-interest and particular loyalties we have. Thus much of our moral deliberation is trying to find ways to marry the person who will want to live
in places where we can help our parents and siblings, or find a job that will allow us to care for our children, or hire another team member who will be able to work well with the other members of our team at work.

3.3.3. Calibration in Extending the Self

Noticing that much of our moral deliberation is how to harmonize among our loyalties leads us to the question of who should be part of our extended self. Here too we find that we have a calibration between moral sentiments and loyalty. Virtue and loyalty are not only calibrated when we are weighting conflicting options of action, but also in deciding our moral framework. The interplay between moral approval and loyalty confines the influence of the moral sentiments to actions on behalf of those associated with the self, while at the same time allowing us to morally reevaluate the limits of the self. How does this work?

While my extended self shapes my obligations, the limits and content of the self are not set. The motives of desires, love and moral approval motivate us to extend the self, while the motives of aversion, hate and moral disapproval motivate us to contract the self. Thus if we desire another person, if we love another person, or if we morally approve of another person, we may be motivated to act on their behalf or take their side. This sort of action, if it is maintained over time, will associate this other person with the self and will give rise to the motive of loyalty. Similarly, if we find that we hate or find morally reprehensible someone who is part of the self, this may lead us over time to disassociate from them and to undermine the motive of loyalty that has so far motivated actions on their behalf.
It is this extension that takes place when a person dedicates his life to caring for the poor or to raising funds for starving people on the other side of the world. These kinds of actions are often considered altruistic, or as motivated by beneficence alone. What actually happens with a person who acts this way consistently, rather than on a whim, is that this person’s moral approval for beneficence takes on a particular object, the poor of a certain city or country, and over time this particular object becomes part of his extended self. For this person, then, caring for these poor becomes an obligation and not only a virtuous act of beneficence. But the obligation of this person cannot be claimed to be an obligation of all people to do the same. For others, who have not taken this particular group or project on, these forms of beneficence are not obligations but merely virtuous actions.

This calibration allows us to, on the one hand, view a great good as not obligatory for us. Thus there could be a great good in curing cancer, and there may be a great virtue in striving for this goal, but I (who have no qualifications to conduct this research) do not have an obligation to pursue it. It would be a violation of my obligations to conduct such research rather than doing my job or taking care of my children, even if on the face of it the good to be achieved by curing cancer is much greater than making dinner and reading books to my children. But on the other hand, the calibration also allows us to reevaluate our own self in light of the moral approval that we feel. So moral approval may challenge a person to take up a career of this kind, and this in turn would change her obligations. This calibration of the two axis allows us to limit our judgements of moral obligation to those actions that we are able to pursues over time, while at the same time challenging us to reconsider the limits of our self and to take up new approved
of objects, that over time will be associated with the self and then will give rise to further obligations.

3.3.4. Calibration in Evaluating Loyalties

What is crucial to understand is that when we are debating between two ways to act, in a particular moment, we are not able to decide in the moment what our extended self is and what are obligations are. This is because loyalties depend on the association of ideas, and thus cannot be decided in one moment. In the moment, our obligations are derived from our existing self, and therefore in the moment of decision, our moral obligation is dictated by our moral approval and loyalties. Our obligations are not decided by moral approval and disapproval on their own, because if they did we would morally approve of actions that conflict and this would pull us in different directions, with no consistency in action and no consistency of the self.

But although this is the case most of the time, we do also evaluate those we are loyal to. Here we evaluate their motives and actions, and not only how we act on their behalf. Here we must tread cautiously, because the moral theories that don’t appreciate the role of loyalty in deciding obligations pull us in the direction of thinking that we evaluate the moral character of others all the time, and that this evaluation directly decides what we are obligated to do. What I mean is that we are prejudiced to think that we are constantly putting our friends, countrymen or even family on the examining table, and figuring out whether they are people we morally approve of and want to associate with. Notice that this would completely undermine the role of loyalty as a component of our moral judgment.
The most common question people ask, when I suggest that obligations come from loyalty, is whether we are obligated by relationships to people we morally disapprove of (obligations to societies we disapproved of will be discussed in the next chapter). Now before we get to the extreme cases, let us put on the table the fact that we often do disapprove of those we are associated with in countless ways. This is not so different from how we evaluate ourselves—we can point to all kinds of ways in which we disapprove of ourselves. It is fully consistent with loyalty—and I would even argue, it is required by loyalty—to find ways to improve those who are part of our extended self. Thus it would not be contrary to loyalty to remind your spouse to call his parents, or to encourage your friends to give charity, or to punish your children when they lie, or to criticize a colleague when she embarrasses another colleague. All the many ways in which we often are in need of improvement are the ways in which we call on each other and help each other improve. In all of these instances, moral approval and disapproval act together with loyalty, rather than in conflict. The tendency to view moral approval and disapproval and loyalty as conflicting with one another is often viewing our role in our relationships as a passive one, when in reality we are able to influence the moral characters of those we are associated with.\textsuperscript{234}

Moreover, loyalty is crucial in these instances, because if we did not have loyalty we would end any relationship with a person who needed some kind of moral improvement, which would be all people. Loyalty is the crucial motive in these situations, because it motivates us to prioritize improving the person rather than walking away from the person. Recall the above

\textsuperscript{234} See for example Ross’s claim that it is our obligation to improve our own character, but does not mention that it is our duty to improve others. Ross, Right and the Good, 21.
example in which Lily’s loyalty to her husband helps her remain faithful to him in the face of her own strong desire to find comfort in the attention of another man. When characterizing obligation we focused on conflicts between loyalty and self-interest, in order to highlight the role of loyalty in dictating obligations. But now we can also ask how does Lily’s moral approval or disapproval of her husband shape her association with him. If moral approval and disapproval were to dictate the contours of the self, Lily’s moral disapproval or hatred of her husband would motivate her to tear up the marriage. But loyalty motivates Lily to remain faithful to the marriage. She could remain faithful to the marriage while hating her husband, but this would be difficult to maintain over time. Thus loyalty in turn requires her to find other ways in which she can give expression to her hatred or moral disapproval of her husband. She is forced to confront him, either with anger or with constructive criticism.

Thus we must be mindful, when we discuss the moral evaluation of those we are loyal to, that this moral evaluation cannot mean that we are only loyal to those that we fully morally approve of. This would undermine the role of loyalty in dictating our obligations, and would render our relationships non-existent. It would also undermine a central component of joint moral activity, which is made possible by loyalty: constant self-improvement and constant joint improvement. Most of the time, moral evaluation has a role in shaping the way we act with regard to our loyalties, without deciding that we must sever them.

This also helps make sense of those close relationships that are so much a part of our extended self, which we do not think they can be completely severed no matter how much we morally disapprove of those we are in a relationship with. Thus we may think that our obligations to our children or parents are so integral to who we are, and are so important to
acting as extended selves, that we do not stop having obligations to them no matter what. It is important to note here that our tendency to see from close up the moral failings of those who are so close to us makes us hyper-aware of their moral failings, especially since we are most likely to be personally effected by them. This is why we are even more careful to not sever the associations that enable these relationships, and thus we are more careful to act on the motive of loyalty to respond to the needs of these who are most closely associated with us.

At the same time, moral approval and disapproval do have a role in choosing, to the extent that we can choose, who is part of our extended self. We are all sure that we do this, but the question to ask is how we do this: how do we evaluate who we include in our extended self, without undermining the very possibility of a consistent extended self?

In order to give some weight to our moral evaluation of others in deciding how central they are to our self, while at the same time not allowing it to get rid of loyalty altogether, we must calibrate the two. Extreme examples may be helpful in making this point, although they are not the way we calibrate the two axis most of the time. Troy Jollimore brings the example of a father whose daughter confides in him that she has killed a person while drunk driving.235 The father who calibrates loyalty with moral disapproval may decide that he must turn his daughter in to the police (I would argue that it is loyalty to his society that does much of the motivation here, not moral disapproval on its own). This is clearly not in the interest of his daughter, and thus is not the most loyal way to act on her behalf. But this disloyal action he performs does not need to undermine his relationship to his daughter completely and is thus not a betrayal. The fact that she

---

235 Jollimore, On Loyalty, 37-46, discussing Andre Dubus’s “A Father’s Story”, in The Times Are Never So Bad, David R. Godine, Boston, MA, 1983. The father in the story does not expose his daughter, but he thinks his decision was immoral.
did kill a person does not mean that she will no longer be part of his extended self. A loyal father will, with great difficulty, continue visiting his daughter and seeking ways to improve her life, while she is in jail. With only moral approval and disapproval dictating moral requirements, there would be no moral obligation for the father to act in this way. But loyalty and moral disapproval are calibrated, so that it is the obligation of the father to keep caring for his daughter.

This example, for all its dramatic effect, is actually an easy case, because the immoral action on the part of the daughter is a one-time action. Although it may have been the effect of an ongoing negligence of human life, in the form of constant drunk-driving, this action does not pull the father into participating in something that he morally disapproves of. This is precisely why the father can, to some extent, separate caring for his daughter from the death that she caused. This is not the case when we are in a relationship with someone who is immoral in an ongoing way. Here we find that since the agent doesn’t want to contribute to immoral actions, over time the other person will in fact become less central to who she is. The change will not be a one-time thing. What is happening, when we disapprove of a relationship and then claim that there is no obligation to support it, is that we are calibrating moral approval with loyalty. It is not the case that loyalty has no role in deciding what are moral obligation is in such a case, but rather that we may over time remove our obligations to this other person by removing them from the extended self. But this means that so long as this person is part of our extended self, there will be obligations that arise from loyalty to him or her.
Conclusion

In this chapter I argued that loyalty is the motive that solves Hume’s problem of obligation, and that Humean moral judgment should be understood as made up of two components, moral sentiments and loyalty.

Starting out with the reading that Hume is a judgment internalist about obligation, I argued that on a Humean picture we must find a motive that can motivate and give rise to the idea of obligation. I then critiqued the best account of Hume’s view that moral obligation arises from the sentiments of moral approval and disapproval. I argued that this account of moral obligation is not satisfactory both because Hume himself thinks it is a weak motivation to action, and because empirically we find that obligations and moral approval do not completely overlap. Applying Hume’s experimental method to the idea of obligation, I argued that obligation can be explained by the causes of loyalty: obligations are actions that fulfill the needs and interests of those who are part of the extended self.

This explanation of obligation lead me to reevaluate Hume’s account of moral judgment. I first argued that my explanation of obligation allows us to have a unified idea of obligation. I then suggested that obligation has a particular function, which is to prioritize among our moral evaluations so that we can act upon them. This prioritization, I argued, is achieved by the prioritizing psychology of loyalty. Finally I argued that moral judgment has a particular form: it calibrates agent-neutral moral sentiments and agent-relative loyalty. This calibration happens on different levels. We both calibrate moral sentiments and loyalty in deciding what actions to takes, and we calibrate our moral sentiments and loyalty when deciding how to extend the self or contract it.
The view I argued for, that loyalty is the source of moral obligation, may be thought to imply that we only have personal obligation, and that we have no impartial obligations. This will be the topic of the next chapter, where I consider moral obligations as they arise from Hume’s artificial virtues.
Chapter 4: Group Loyalty and the Psychology of Impersonal Obligations

In chapter 3 I argued that loyalty is the reflective impression that gives rise to the idea of obligation, and thus is the psychological source of obligation. But the last chapter left open the scope of actions which loyalty can motivate, and thus the scope of obligation. More specifically, we saw that loyalty motivates us to act on behalf of those who are in our “close circle”, but this left open the question whether loyalty motivates us to fulfill impersonal obligations. We clearly think we have impartial obligations to people who we do not know, and it is not clear how the explanation I have given of obligation can account for such obligations.

The move from personal to impersonal moral judgments plays a central role in Hume’s moral theory, distinguishing the natural and artificial virtues. In developing the category of the artificial virtues, Hume takes the controversial position that impersonal obligations such as respecting private property and keeping promises do not arise from intuition or a priori rules, but rather arise from convention. In this chapter I argue that group loyalty can extend to those with whom we share a convention, and thereby motivate us to act according to the convention. If loyalty can motivate us to act according to Humean conventions, then it can motivate us to return loans, keep promises, and other such impersonal obligation. This will not mean that loyalty can motivate us to act in all ways we approve of, but will greatly extend the category of obligations explained by the causes of loyalty.

My category of personal obligations and Hume’s category of the natural virtues do not completely overlap. On Hume’s account of the natural virtues, we approve and disapprove of the motives that lead us to act in certain ways to strangers, in individual instances. Yet Hume’s view that we primarily judge a person’s character according to that person’s treatment of those in his close circle led me to conclude that personal obligations arise in our relationships with those in our close circle.
In order to make this argument, in part 1 of this chapter I argue that Humean loyalty takes groups, but not rules, as its object, meaning we can be loyal to groups but not to rules. In part 2 I challenge the view that impersonal obligations arise from impartial rules. I show that on Hume’s account of convention it is partiality towards a group that leads to adopt a convention and morally approval of it. For Hume this is an epistemological claim, since Hume thinks we only have the evidence necessary to judge that a rule is moral once it is adopted as a convention by a particular society. In part 3 I argue that loyalty is the natural motive to act in accordance with the convention, once it is up and running. In part 4 I offer my account of impersonal obligations: I argue that idea of impersonal obligation arises from the reflective impression of loyalty that is caused in instances where self-interest motivates us to act against the convention. In the last section I discuss some normative implications of my explanation of moral obligation.

1. Recognizing Group Loyalty by Its Causes

In my analysis of loyalty I argued that loyalty depends on an association of another person with the self. This kind of association takes time and constant conjunction. In the case of impersonal obligations, the beneficiary of the obligatory act has no previous relationship to the agent obligated. I have to respect the property of others, whether I have had any previous interaction with them or not. Thus it seems that the psychological requirements for loyalty just don’t exist in these cases. There is no association of the recipient with the self, and therefore there can be no loyalty that motivates me to fulfill my obligation.
Before we address this question, and show that loyalty can explain impersonal obligations, we must extend the view of loyalty we’ve been working with. Until now I focused on personal obligations, and therefore I focused on loyalty to individuals. In this part I argue that we have the same motive of loyalty not only to individuals, but also to groups, institutions and nations, but cannot be extended to rules or principles. We recognize the motive of loyalty by its causes, which are the needs of those who are part of the extended self. While groups have such needs, rules do not.

How can we be loyal to a group of people? We clearly cannot associate every person in the group with our self, since we do not have the constant conjunction or causation necessary for association. Theories of loyalty that define it as a strong emotional connection may find the claim that we can be loyal to a group difficult. It seems that you don’t have the same care and concern, or love and affection, towards every person in a group of people as you do towards a spouse or child. More importantly, there is an intimacy in close relationships that cannot be felt towards a group. Is it possible that we have the very same motive to act on the interests of our family members, and to act on the interests of our university, city or nation?

My theory of Humean loyalty did not characterize loyalty as a strong emotional tie to a person. Rather it showed that loyalty is a strong motive, because it is caused by the needs of someone who is associated with the self. On Hume’s view of the self as a bundle of ideas,

---

237 For such a view of loyalty as a strong emotion, see Simon Keller, Limits of Loyalty, 3. Kleinig undermines this claim, by arguing that emotions are not as central to loyalty as supposed, although they often accompany loyal action. He points out a standard example of loyalty, which does not seem to require a strong emotional tie. If a wife who no longer loves her husband resists the desire to have an affair, because of her vow to her husband, she is being loyal although she does not have a strong emotional attachment. See Kleinig, On Loyalty and Loyalties, 15-17.
another person can be one of those ideas that make up the self. This extension of the self to include other people brings about the motive to act on behalf of these people.

The association of ideas that makes up the self does not require that the idea associated with the self be a particular person. Thus the motive of loyalty does not require that the object of loyalty be an individual person. Rather the association of ideas requires that the idea be particular, and therefore loyalty requires that the object be particular. What this means is that we can be loyal to any object that is a particular entity, not only to individual people. A group of people is a particular entity, and therefore it can be associated with the self in the same way that an individual can be associated with the self. One can have a past history with a group, just as one can have a past history with an individual. One can be proud or humiliated by a group, just as one is proud of humiliated by an individual.

Particular individuals and groups can be distinguished from traits, qualities, and rules which are not particular and cannot be the objects of loyalty. Recall that the cause of loyalty is the need of someone who is part of the agent’s extended self. Thus the objects of loyalty must be entities that have needs, and can be strengthened or weakened. If a group has its own needs, then it can be the kind of object we can be loyal to. Unlike a group, that has needs, a rule has no needs. A rule is more like a quality or a category than a particular entity. Since rules doesn’t have needs, and cannot be strengthened or weakened, rules are not something we can be loyal to. A tradition, on the other hand, is something particular – it can die out, as the tradition of the Aztecs has died out, and it can be strengthened, as the tradition of scientific inquiry was strengthened by

---

238 For the view that loyalty’s object is always a particular entity see Oldenquist, “Loyalties”, 175; Keller, Limits of Loyalty, 17; Alasdair MacIntyre, “Is Patriotism a Virtue?”, The Lindley Lecture, University of Kansas, March 26, 1984, 4.
the discoveries of Newton. Thus a person can be loyal to a tradition, but cannot be loyal to a rule or a principle.

In the previous chapters I argued that the causes of loyalty are the needs of the person associated with the self. It is important for our discussion to distinguish and show that groups have their own needs, which are not reduced to the needs of all the individuals that make up the group. Thus when a person responds with loyalty to a group, they are responding to the need of the group, rather than the needs of individuals that make up the group.

Let us take an example to make this distinction clear. When a person is part of a football team, they associate the team with their self. The team is a particular entity, with its own history, goals, successes and failures. These can be distinguished (to some extent) from the individuals who make up the team. Team members may change, but the team remains the same as long as it has this ongoing story. Each member of the team (or other people, for that matter) can come to have the team associated with themselves. They have a history of playing with the team, which creates this association. This association can extend the self of the players to include the team, and this extension of the self will offer them a motive of loyalty to act on behalf of the team. For example, a team cannot win without the contribution of its members, so the team members will be motivated to play hard, even when this conflicts with their own self-interest.

The team is a small society, in which every team member still knows all the other individuals involved. Thus it may not seem at first that this is a helpful comparison, if we are interested in loyalty to a large society. But what is important to notice here is that although a player could be loyal to each and every member of his team, loyalty to the team is distinguished from loyalty to each individual on it. We can distinguish loyalty to the team from loyalty to the
individuals who make up the team. The team is associated with the self of the team member as a particular entity, in addition or alongside loyalty to any individuals who are part of the team. Just as the team can be associated with the self, because it is a particular entity with a particular history, a much larger group of people that makes up a particular entity can also be associated with the self, so long as it is particular.

To conclude, although we don’t have the same intimate feelings towards groups as we do towards individuals, we do have loyalty to groups. Groups are particular entities, which can be associated with the self, as are institutions, nations, and traditions. These entities have needs, and succeed and fail. When their need or interest becomes apparent to a person who has them as an idea that is part of his or her self, they cause the reflective impression of loyalty. As opposed to groups, rules and principles are not entities that we can be loyal to. Although we may be able to associate rules and principles with the self, on Hume’s account of association, rules do not have needs, and thus cannot cause the motive of loyalty.

This explanation of how it is that we can be loyal to groups underscores the problem with the possibility that loyalty is the impression that gives rise to the idea of non-personal obligations. Non-personal obligations are obligations that arise independently of the characteristics of the recipient. But loyalty is precisely a reaction to a particular object’s association with the self. The very particularity of the object of loyalty, makes it seem to be the wrong kind of motive for acting on non-personal obligations. This is the question I turn to in the next part.
2. Challenging the Impartiality of Impersonal Obligations

As discussed in chapter 1, Hume thinks that we have two kinds of virtue, the natural and the artificial. In the first chapter of the dissertation we mostly focused on the question what is the difference in judging that the natural and the artificial virtues are virtues. We showed that in the case of the natural virtues we have moral approval of a particular motive in every instance that it appears, such as approval of beneficence or courage. Thus the natural virtues do not require any societal context, in order to judge that they are virtues. In the case of Hume’s artificial virtues, such as justice or honesty, it is only once there is a convention in place, which people generally follow, that we approve of the motive that give rise to these virtues.

What I want to focus on here is that Hume’s distinction between the natural and the artificial virtues is a move from virtues pertaining to personal relationships to those pertaining to interactions between people who do not know each other but live in the same society. As Mikko Tolonen and others have emphasized, Hume’s discussion of the artificial virtues is the discussion of virtues that come about in large societies, where we don’t know everyone.\textsuperscript{239} It is therefore in the artificial virtues that we must look for Hume’s account of impersonal obligations.

We take it for granted that keeping a promise or returning a loan are impersonal obligations. Hume offers a revolutionary account, on which the impersonal obligations of keeping one’s contract, fulfilling promises, and other impersonal obligations are only morally approved of in the context of conventions adopted by a certain society.\textsuperscript{240} My claim will be that although conventions are impartial rules, there is a partiality to one’s society that underlies the

\textsuperscript{239} For the explanation of the movie from natural to artificial virtues as the move from small to large society, see Tolonen, \textit{Mandeville and Hume}, 167, 174, 181; Forbes, \textit{Hume’s Philosophical Politics}, 75.

\textsuperscript{240} I focus, for brevity’s sake, on Hume’s account of justice. But the story is similar in the other artificial virtues.
adoption of Humean conventions and the approval of it. I thus oppose the view that conventions are a priori moral rules, and argue that they are tools to benefit a particular society. , while the rule is a tool applied in order fulfill our obligation that arises from our loyalty to the group. If this account is correct, I argue, then Hume presents impersonal obligations as arising from loyalty.

2.1 Hume on Limiting Partiality

Before we consider my interpretation that there is group loyalty at the basis of convention, we must point out that Hume thinks convention is adopted to curb our partiality. Hume says that partiality towards one’s own is the tendency that we limit, when we move from the small to the large society. Although others think that it is selfishness that prevents human beings from living in society, Hume is convinced that it is not selfishness but partiality towards those close to us. He supports this claim by arguing that human beings are not quite as selfish as they are sometimes made out to be:

    So far from thinking, that men have no affection for any thing beyond themselves, I am of opinion, that tho’ it be rare to meet with one, who loves any single person better than himself; yet ’tis as rare to meet with one, in whom all the kind affections, taken together, do not over-balance all the selfish…

Although Hume thinks human beings are not selfish, he thinks they are decidedly partial, and it is this partiality which he thinks needs to be overcome for human beings to live in large societies. People are naturally inclined to prefer those who are close to them over strangers. Thus even if

---

241 T 3.2.2.5, 486-7. “For it must be confess, that however the circumstances of human nature may render an union necessary… there are other particulars in our natural temper, and in our outward circumstances, which are very incommodious, and are even contrary to the requisite conjunction. Among the former, we may justly esteem our selfishness to be the most considerable.”
narrow self-interest does not prevent us from living in society, because it is limited by our concern for those close to us, this second tendency itself prevents us from living in the larger societies that are necessary for human flourishing.

But tho' this generosity must be acknowledg'd to the honour of human nature, we may at the same time remark, that so noble an affection, instead of fitting men for large societies, is almost as contrary to them, as the most narrow selfishness. For while each person loves himself better than any other single person, and in his love to others bears the greatest affection to his relations and acquaintance, this must necessarily produce an opposition of passions, and a consequent opposition of actions; which cannot but be dangerous to the new-establish'd union.242

Hume thinks that even if we are not selfish, our tendency to prefer those who are close to us will conflict with this same tendency in others, and will prevent us from uniting in a larger society. Thus according to Hume the very existence of prosperous societies is dependent on limiting our partiality towards those who are our own.

Hume thinks that our partiality does not only motivate us to act in ways that undermine society, but also influences our judgment. When people are motivated by partiality towards those who are close to them, with no limitation on this partiality, they will also approve of such actions when no conventions are in place.

Now it appears, that in the original frame of our mind, our strongest attention is confin'd to ourselves; our next is extended to our relations and acquaintance; and 'tis only the weakest which reaches to strangers and indifferent persons. 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… our natural uncultivated ideas of morality, instead of providing a remedy for the partiality of our affections, do rather conform themselves to that partiality, and give it an additional force and influence.243

242 T 3.2.2.6, 487.
243 T 3.2.2.8, 488-9.
Hume explains here that our strong propensity to prioritize the needs of those who are close to us over those of others, our natural virtues will lead us to approve of people who act partially. In order to explain how we come to judge that it is virtuous to keep a promise to a stranger or repaying a loan to a person we never met, even if it comes at the expense of our close relations, Hume thinks that human beings must place limitations on their partiality. These limitations are put in place, I shall argue, not by transcending partiality but rather by taking on larger group loyalties.

2.2. Partiality in the Adoption and Moral Approval of the Convention

In this section I argue that although Hume thinks that conventions come to limit the partiality towards our close friends and family, he can be read as arguing that what limits partiality is not an abstract rule of conduct but a wider loyalty. Although we approve of a convention, which dictates impartial behavior, it is loyalty towards the group that motivates joining the convention brings about the moral approval of it. Thus on my interpretation of Hume’s account of convention, impartial rules are tools that tell us how to act when we wish to achieve prosperity and security for a particular group of people we are loyal to, rather than a priori principles that dictate obligations independently of partial ties.

Richard Rorty has argued that the virtue of justice should be understood as a form of the virtue of loyalty. Rorty raises the question: “Would it be a good idea to treat the word ‘justice’ as the name for loyalty to a certain very large group – the name for our largest current
loyalty, rather than the name of a virtue distinct from loyalty?“ Rorty is aware that justice is usually described as an impartiality, and is opposed with loyalty. But he continues “The question is: should we describe such moral dilemmas as conflicts between loyalty and justice or, as I might suggest, as conflicts between loyalty to smaller groups and loyalty to larger groups.” Rorty argues that the rules of justice, that dictate non-preferential treatment, are not abstract rules that apply regardless of any loyalties. He argues that “Plato and Kant were misled by the fact that abstract principles are designed to trump parochial loyalties into thinking that the principles are somehow prior to the loyalties.”

As was discussed in chapter 1, Hume thinks that there is no natural motive to follow the rules of justice, where by justice he means rules of private property. Thus there is no natural motive to return a loan to a stranger, to respect the ownership of property by others, or to demand the return of one’s own property. There is also no natural virtue to follow the rules of justice, meaning there is no virtue (and no obligation) to respect private property when one does not live in a society where there is such a common way of acting. Here Hume is clearly in opposition to John Locke, who argued that the right of private property is self-evident to every reasonable person prior to the existence of society. He is also disagreeing with any view that holds that private property is moral rule that is known a priori. This general position Hume holds in his account of justice leads me to suspect that he shares Rorty’s view that justice is taken up in order to fulfill a loyalty.

245 Ibid, 140.
246 Ibid.
As we just saw above, Hume clearly thinks that the rules of justice come to limit our partiality towards those who are close to us, what Rorty calls “parochial loyalties”. What I wish to argue is that although Hume thinks that our unlimited loyalty to our friends and family would lead to constant conflict with others, he thinks that we do not overcome this conflict by adopting preexisting rules, but rather by taking up a loyalty to a larger group of people with whom we conduct the epistemological enquiry that leads us to approve of the convention of justice. Thus Hume agrees with Rorty for epistemological reasons. For Hume, the way we come to know that acting honestly is a virtue, is by living in a particular society in which we can see the benefits of general compliance with the convention. Only once the members of the group are loyal to one another, can they have the assurance of cooperation that is necessary for approving of the convention. In order to make this claim, we must discuss the role partiality plays in Hume’s account of justice.

Recall Hume distinguishes “the manner, in which the rules of justice are establish’d”, and the “the reasons, which determine us to attribute to the observance or neglect of these rules a moral beauty and deformity.” First, he gives a genealogical account of how it is that a society comes to have a convention of private property. Then he offers an account of why we approve of the motive to act according to the convention, thereby rendering it a virtue. In the following I show the need for partiality in each of these stages.

248 T 3.2.2.1, 484.
2.2.1 Partiality in Adopting the Convention

Let us first turn to Hume’s genealogical account. In this passage Hume describes the circumstances that lead to the adoption of the convention of private property:

This convention is not of the nature of a promise: For even promises themselves, as we shall see afterwards, arise from human conventions. It 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; since the actions of each of us have a reference to those of the other, and are perform’d upon the supposition, that something is to be perform’d on the other part.249

Hume claims that there is “a general sense of common interest” which “all members of the society express to one another” and it is this common interest that is voiced and acknowledged which “induces them to regulate their conduct by certain rules”. Thus the convention taken up is not an abstract rule that is good or required for any society or any human being to follow. Rather the convention is taken up as way to regulate conduct in the particular situation that this particular group is in. The situation of having “a general sense of common interest”, which they express to each other, causes the group to take up the convention. Thus the requirements of the convention are contingent on the nature of the group of people in question, and the way they interact with one another.

Now Hume is not very clear on what this “common sense of interest” is. When further explaining the situation, Hume says that “I observe, that it will be for my interest to leave

---

249 T 3.2.2.10, 490.
another in the possession of his goods, *provided* he will act in the same manner with regard to me.”\(^{250}\) Thus the common reading of this stage of the story is that every person joins the convention out of self-interest. But recall that for Hume, self-interest encompasses one’s close friends and family, so it is actually what Rorty called a “parochial loyalty”. I thus suggest that what motivates people to join the convention is loyalty to those who are their close associations.

Whether it is self-interest or parochial loyalty that initially motivates us to join the convention, Gerald Postema has argued that the point in the story where Hume thinks that a group of people take up a convention of private property is a moment when the people in question take up a common point a view, a point of view from which they ask what is good for this group they are part of. He says the following:

> At this point in his story, Hume speaks of “a common sense of interest” (*Treatise*, 490) and “a sense of interest suppos’d common to all” (*Treatise*, 498), rather than “a sense of common interest”. The language marks an important difference, for what emerges is not merely a sense that there is a coincidence of interest, or even a new interest in some common good or goal, but a *new practical attitude*, a shift in perspective on the practical problem facing the parties… The shift does not by itself solve or dissolve the practical problem facing the parties; rather, it transforms it. Conflict among the parties may still remain, but the parties no longer view their conflict from their respective private or individual perspectives, but rather from a common one.\(^{251}\)

On Postema’s reading, the individuals may be motivated by self-interest to join with other people, but once they have decided to join with others, the practical perspective from which they judge is the practical point of view of what is beneficial to us as a group. Postema is not denying that Hume thinks self-interest motivates us to join with others at first, but he is claiming that

\(^{250}\) T 3.2.2.10, 490.
even if this is the case, at some point a joint effort requires a shift in perspective. This shift in perspective does not eliminate one’s ability to continue to judge whether the actions that are good for the large group are also good for one’s narrow interest. But it does allow us to make another calculus, along with the private one, what is good for the group as a whole.

Postema only talks about a perspective, a point of view from which one judges what to do. He does not offer a motive that is different from self-interest. Postema emphasizes the plural perspective from which we take up the conventions of justice, according to Hume. But I would like to emphasize the partiality involved. If Postema is right, and we take up the convention from a plural perspective of “we”, then this is precisely the particular group that we can view as our own and be loyal to. Thus the perspective that Postema is describing is an epistemological one, but it goes together with the possibility of a motive of loyalty. His description fits very closely with the description I offered of the perspective of the loyal person. We both described a point of view from which we ask what is good for the “we”, rather than for the individual alone (recall the husband who is deciding whether to move for his wife’s dream job). What my theory of Humean loyalty adds to Postema’s description is the motive of the person who takes up the first person plural perspective. The motive to ask what will enable this group of people I am a part of to live together is the motive of loyalty. If the first person plural perspective Postema describes is merely a perspective, one could take it up out of sympathy with others or out of self-interested calculation. But if one is to act upon this perspective, one needs to be motivated by loyalty.

But notice that having a common sense of interest is not sufficient for people to join in a convention. The other requirements is “that the actions of each of us have a reference to
those of the other, and are perform'd upon the supposition, that something is to be perform'd on
the other part.” Thus it is not self-interest alone that causes the formation of a convention, but
rather self-interest together with an assurance of cooperation. What allows for this kind of
assurance that others will cooperate?

I suggest that the assurance that others will act according to the convention, which
Hume suggests is the prerequisite for taking up the convention, is an expression of mutual
loyalty among the members of the group. Of course the members of the group cannot know for
sure how others will act – and clearly a convention doesn’t ensure that everyone will act in a
certain way. But the common feeling Hume is describing can be felt before taking up the
convention. What the members can see is that the group has become its own entity, associated
with the self for each of the members. The way this can be tested is the way people react
emotionally to the group as a whole. If others feel pride and humility, and other emotions, caused
by the success or failure of the group, then we can know that the members have associated the
group with the self. If this is the case, then the motive of loyalty is available to these members,
and can act as an assurance (to the extent that there is ever an assurance) that others will act upon
the convention.

Would Hume have agreed that people joining a convention share a motive of loyalty
prior to taking up the convention? As I said, Hume is being unclear here. On the one hand, he
says that the motive to join the convention is self-interest. On the other hand, he talks about a
common interest which is understood by the individuals which is a prerequisite to forming the
convention. The confusion here is one that often takes place, when we choose to adopt a loyalty
that is broader than the one that we had before. On the one hand, one is motivated by the existing loyalty to one’s close circle of family and friends to look for cooperation with others that will bring peace and prosperity to those one is already loyal to. One does not take up the project of coming to terms with other people about cooperation, without ensuring that this benefits one’s existing associates. As Hume says regarding those “parochial loyalties”:

Nor is such a restraint contrary to these passions; for if so, it cou'd never be enter'd into, nor maintain'd; but it is only contrary to their heedless and impetuous movement. Instead of departing from our own interest, or from that of our nearest friends, by abstaining from the possessions of others, we cannot better consult both these interests, than by such a convention; because it is by that means we maintain society, which is so necessary to their well-being and subsistence, as well as to our own.

If loyalty to those in our close circle does lead us to cooperation with others, this cooperation may lead to the creation of a larger, particular, group that brings about another loyalty. At some point, we are able to view the larger group created, which becomes its own particular entity, as one that calls on our loyalty.

Although Hume is unclear on the particular motive, he is quite clear that the convention is taken up as a solution to the problem of a large group of people. Thus he is quite clear that the rule is subsidiary to the relationship among this group of people. Just as the common point of view for the natural virtues was the point of view from which we sympathize with the real people surrounding a particular person, whose character we judge, thus also in the

---

253 T 3.2.2.9, 489.
artificial virtues the common sense of interest is born out of partiality towards this particular group.

2.2.2 Partiality as Prerequisite to the Moral Approval of Convention

So far we’ve discussed the part of Hume’s story that tells how we come to adopt a convention. Hume thinks this is a first stage, followed by another stage in which we judge that the convention is good. Here, too, partiality towards the group is required. Recall that what causes us to approve of the convention is its benefits to society:

Tho’ justice be artificial, the sense of its morality is natural. ‘Tis the combination of men, in a system of conduct, which renders any act of justice beneficial to society. But when once it has this tendency, we naturally approve of it; and if we did not so, ‘tis impossible any combination or convention cou’d ever produce that sentiment.²⁵⁴

Sayre-McCord emphasizes that Hume thinks conventions are rules that are beneficial to all of their adherents, although they might not benefit them to the same extent.²⁵⁵ “For Hume, the conventions that underwrite artificial virtues are restricted to those that are mutually advantageous,”²⁵⁶ says Sayre-McCord. Conventions such as fashion, that are not clearly advantageous to all, are not conventions and do not give rise to any virtuous motive.

But it is not enough for a rule to be mutually advantageous for it to give rise to an artificial virtue. Sayre-McCord explains that the possibility that a rule will be advantageous is not sufficient to make it a convention. A rule must be accepted by a group of people for it to be a convention: “Significantly, on Hume’s view, it matters that the conventions in question actually

---

²⁵⁴ T 3.3.6.4, 619.
²⁵⁵ There could be different advantages to different people. See Sayre-McCord, “Hume on the Artificial Virtues”, 448.
²⁵⁶ Ibid, 447.
be in place. That a convention would be advantageous, if only we had it, may be reason to work to establish it, but it is no reason to act according to it absent the participation of others.”

This is quite different from how we usually think of morality, where we think that what is moral is not contingent on what people in fact do.

Hume explains why it is that a rule must not only be theoretically good, but accepted, for it to be morally approved. This is because in order to approve of the convention “’tis requisite not only that it be advantageous, but also that men be sensible of its advantages.” But the advantages of the convention cannot be known through speculation: “’tis impossible, in their wild and uncultivated state, that by study and reflexion alone, they should ever be able to attain this knowledge.” A person who is not raised in a society with a convention, has no way of knowing that a convention would be beneficial. It is only through our experience of people following the convention that we know it is good. Convention “arises slowly, and acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it.” Like the usefulness of language or money, the usefulness of the convention is learnt through experience. This is because the advantages become known through experience of what it is like for people to act according to the rule.

Why can’t the advantages be known through speculation, and the rule be a convention because it would be advantageous? It seems to me that this goes back to people’s motivation, and Hume’s claim that conventions only arise when people can rely on each other.

---

258 T 3.2.2.4, 486.
259 Ibid.
260 T 3.2.2.10, 490.
261 Hume’s comparison to convention in T 3.2.2.10, 490.
Hume thinks that the actions of a convention are the kinds of actions that are only good through common compliance. What makes the system good is everyone acting according to the rule and the good consequences are a product of this general compliance. So it is not so simple to know what the consequences of these actions would be like, without trying it. Hume’s view is that we only find following the convention virtuous once it is up and running, and we see what its consequences are over time.

But no less important, is that speculation cannot tell us what rule will in fact garner the support of people. It is through the actions of a certain group of people that the rule becomes a convention. A convention is always a rule that belongs to a certain group of people, who have turned it into a convention by complying with it. Since the convention depends on compliance, it is only their compliance that turns the rule into a convention. The known advantages become known through their actions, so that the advantages to them of the convention are only known through their actions.

The convention is not any good rule, it is a rule that has, through the adoption of a group of people, puts limits on those people’s choices and actions. They are not only motivated by the advantages of the rule, since other rules may have advantages too. As Sayre-McCord puts it, the existing conventions limit our options of acting according to what could be advantageous: “once we find ourselves within a mutually beneficial conventions, the range of available options is constrained by our established obligations.” This is because there could be any number of rules that might be advantageous, but only the one accepted by the people following it is a

---

262 Baier points out that Hume is talking about property laws, and is likely thinking of the complex Scottish system, where every small part of the law does not seem to be good in any way. Baier, Cautious Jealous Virtue, 67.
263 The known good is desired more than an unknown good. See T 2.3.6.2-5, 424-426.
convention. Conventions constrain our actions, giving us a reason to act by limiting our options from all the ones that might be good to the ones that obligate. Hume is pointing to the fact that when we adopt a convention, we choose a specific set of advantages that are brought about through our cooperation with a specific set of people. Since any number of advantages could be achieved through rules, the only way to achieve the advantages of a certain rule is by being loyal to a group of people. Thus loyalty is the missing link which explains how the relation of convention to a group of people who shares it influences the will and defines obligations.

Hume thinks that conventions are adopted through an epistemological process of a particular group of people. As Postema argued, individual agents choose to take up the convention of justice by considering whether it is good for “us”, a group that they are a part of. Moreover, the way we judge that the convention adopted is a moral one is by considering its consequences for the group as a whole. The particularity of the group in question makes it possible to be loyal to a group that has taken up the convention. Just as “Virtue and vice, when consid’d in the abstract… excite no degree of love or hatred”, so too loyalty is not to the rule but to the group who comes together around the rule.

A similar view is voiced by Oldenquist in his discussion of loyalty:

A loyalty defines a moral community in terms of a conception of a common good and a special commitment to the members of the group who share this good. The members, along with certain conventional, institutional structures, and often a geographical location, together constitute the community that is the object of my loyalty. Those who share this common good comprise my tribe; the common good is its flourishing, and this is why we acknowledge a system of social morality whose purpose is the safety and flourishing of the tribe and which applies impartially to its members.266

265 T 2.2.1.7, 331.
266 Oldenquist, “Loyalties”, 177
Rorty and Oldenquist do not think that this is a relativist view, in the sense that there is no way to adjudicate between societies that have rules of justice and those that do not. Rorty thinks that if we compare societies that have gotten rid of slavery to those that have not, we will clearly think that those that have gotten rid of slavery are better. But Rorty thinks that the way we come to get rid of slavery is not because we have an abstract rule that preceded any loyalty, that slavery is wrong. Rather we search for the best rules to live by in the society that we live, out of loyalty to this society, and this leads us to conclude that slavery should be removed.

At times, it sounds like Hume thinks that experience everywhere will teach human beings the benefits of private property. But the spreading of a convention to other societies is not what makes private property a convention. It is not constitutive of the convention that it exist in other societies. All the convention needs is to be accepted among a particular group of people. This is why Hume emphasizes the times and places in which property is not relevant. The fact that property is not relevant in small or violent societies, doesn’t undermine approving of it in places that do have such a convention and find it mutually advantageous. Hume’s comparison to language and money is apt: although these exist in many human societies, the usefulness of a language or a coin in one society doesn’t depend on their existence in another society.

2.3 Rejection of the Interpretation that Conventional Obligation Comes From Rules

So far we’ve seen that the approval of justice comes from its beneficial consequences for a particular group of people. The impartiality of justice is seen as good to the extent that it

---

268 T 3.2.2.14, 492-493.
has beneficial consequences for all members of the convention, not because impartiality is moral or good in itself.

Now that we have seen that there is partiality required for adopting a convention and for judging that a convention is good, it is time to discuss obligation. Considering obligations to arise from rules is a common view. The very impartiality of rules is considered the source of obligation by some, while according to others it is our ability to regulate our conduct by rules that is the source of obligations. I will not discuss this theory in general here. Instead, I address an interpretation of Hume’s theory of impersonal obligations as arising from impartial rules.

Stephen Darwall has read Hume as suggesting that moral obligations in the case of the artificial virtues arise out of our ability to live by rules. Darwall agrees with the claim that I made in the previous chapter, that obligations are distinguished from the good. But he thinks that what puts a limit on such a calculus are rules. “The just person does not act simply from desire for the good; she regulates her conduct by rules of justice. In accepting these, she regards action falling under them as what she ought or must do.” 269 The basis Darwall finds for this is that Hume often talks about just persons are ones who “lay themselves under the restraint” of rules 270 and regard the rules of justice as “sacred and inviolable.” 271 A just person, Hume writes, “must fix an inviolable law to himself, never, by any temptation, to be induc’d to violate” the rules of

---

269 Darwall, *British Moralists and the Internal ‘Ought’*, 317. Darwall contends that Hume thought there are three kinds of obligation: natural obligation (motivated by self interest), moral obligation (motivated by moral approval) and rule obligations that are motivated by rule following. Ibid, 296-298. This view is similar to the pluralist views of obligation I mentioned in the last chapter.

270 T 3.2.2.24, 498-500.

271 T 3.2.6.10, 533; T 3.2.10.15, 562-3.
justice. Thus Darwall reads Hume in a Kantian way, where obligation comes from the very ability to define rules for ourselves.

This view of the artificial virtues has some merit, since the artificial virtues do require a code of conduct that becomes a rule in order to be virtues at all. We can’t be virtuous for acting honestly, according to Hume, if there is no rule of acting honestly. Thus on Hume’s own account, rules have a central role in defining impersonal obligations, in a way that they do not in personal relationships. We can take justice as our prime example of this move to the impartiality of rules. When we comply with the convention of justice we put aside many of the differences between people that motivate us, especially how close these people are to us. Thus when we act justly, we view many of the distinctions between people as irrelevant, and we view them according to rules that are blind to our personal partiality towards them.

And yet the fact that impersonal obligations are characterized by rules of conduct does not mean that the psychological source of these obligations is the rule itself. In the following I show that although rules have an important role in judging what our moral obligation is, these rules do not replace loyalty as the motive that gives rise to the idea of obligation. I show that a Humean approach to impersonal obligations is viewing them as dictated by conventions. But these conventions, for Hume, are judged to be good because they contribute to the success of a particular society, and thus conventions are not obligating in and of themselves, but rather are a necessary tool in fulfilling one’s obligation to one’s society. My claim is that the obligating

---

272 T 3.2.2.27, 501; Darwall, *British Moralists and the Internal ‘Ought‘*, 293.

273 Wiggins suggests that we can read the obligations of Hume’s natural virtues as also generating a kind of rule. This is because we cultivate our moral sentiments by a standard that is adopted through discussion with other people. Wiggins, “Categorical Requirements: Kant and Hume on the Idea of Duty”, 83-106. This view will be addressed below.
nature of conventions doesn’t arise from them being rules. Rather the obligation of these conventions comes from the same source of obligation of the natural virtues, namely loyalty. Although the convention does limit our loyalty to those in our close circle, it does so by offering a competing, wider loyalty, that encompasses the loyalty to those in our close circle but widens to include others. This wider loyalty is what puts a limit on our personal loyalties, rather than offering an alternative source of obligation.

What is then the need for rules? Why don’t we just fulfill the needs of the groups we are associated with, with no rules? I offer three uses of rules, that could be relevant here, but do not view the rule as the source of obligation.

Andrew Oldenquist offers a theory, on which abstract rules of justice are the way we choose to interact within the confines of a group we are loyal to. “Our wide and narrow loyalties define moral communities or domains within which we are willing to universalize moral judgments, treat equals equally, protect the common good, and in other ways adopt the familiar machinery of impersonal morality.” But Oldenquist’s view still assumes that there is something valuable in impartiality itself, only that we confine it to communities or domains where we are able to act on it. This still seems far from Hume’s view, where the convention is only adopted as the best way to live together. Why would it be the best way to live together with others, to be impartial?

First, Oldenquist suggests that our rules of justice are not so different from the equal treatment parents give to their children. We can expand this claim, suggesting that a group

---

275 Social morality is impersonal and impartial when it confines itself to intratribal considerations, as when a parent, facing no challenge to family loyalty itself, seeks to do the right thing when interests conflict within his or her family. But when the national good competes with the good of one’s family, the latter appears as a naked loyalty,
stays together best if membership in the group offers each member an equal standing. Within a group that we are loyal to, the members of the group have an equal place, and this is why they receive equal treatment. On this suggestion, we have no use for equality between those who are loyal to and those who are trying to purposefully harm us. But we do have use for impartial treatment among those who wish to live in harmony with us, because this equal treatment motivates the members of the group to prioritize being part of the group. As Oldenquist writes, “According to this scheme every loyalty has the double aspect of being a self-dependent value in intersocietal competition and of grounding a system of impersonal, social morality for adjudicating intrasocietal conflicts.”

The second use of impartial rules is in allowing for coordination. My understanding of Hume is that the use for rules is that they enable coordination of people who share loyalty to one another, but need guidance in what ways to express this loyalty. As opposed to close relationships, where we know the individual well and therefore are conscious of their needs, in larger societies we don’t have a clear idea of what ways we can assist others. Moreover, we are not able to take on the interest of each individual in full. We would not be able to take on all interests of such a large group of people, with no clear limits, as we sometimes do with our close family and friends, because this would be too large a responsibility. The convention is pivotal, because it allows us to share some limited and specific interests with a large group of people, and allows us to be loyal to these people in the particular shared interest we have.

---

and the former as social morality that demands impartial consideration of family and nonfamily. So too, nationalism in turn is a mere loyalty from the wider perspective of utilitarianism, but takes the form of impersonal, social morality in internal matters, including clashes with nested, narrower loyalties such as community or civic loyalty. Oldenquist, “Loyalties”, 177-178.

276 Ibid, 178.
This view of impartial rules as a tool, that points out in what ways they must be loyal to each other, also allows us to extend the group we are a part of so long as new members take up the rule. We often talk about the loyalty to the team as motivation to play against others more ferociously. We don’t often ask what motivates us to follow the rules of the game, which are common to my team and other teams. My argument is that our joint interest to play football unites me, my team members, and the team members of the other competing teams. As strange as this might sound, the loyalty to the team leads me to extend a wider loyalty to all those who play the game with me in the league. My loyalty to members of other teams is definitely not a loyalty to them as individuals – as individuals they are my enemies, members of competing teams. My loyalty to them is a loyalty to members of a larger, mutual group we cooperate in to make it possible to play the game.

The need to play with others, who are competitors, requires setting up clear rules of the game. There is no way to win or lose without these rules. What motivates a player to act according to the rules, when they can cheat and thus their team could more surely win? I suggest that what motivates in such an instance is not mere rule following. The rules are a tool, to ensure that competitors can play and win. The motive to follow the rules of the game is loyalty to the league, since I must be able to extend a larger loyalty to those who play the game with me.

Postema warns us that although rules are necessary for being moral, in larger societies, we should not view the rules themselves as deciding what is moral. the common sense of interest is not itself a rule. Rather, it involves a shared commitment to adopt and follow a set of rules should one be available to the parties, and a perspective from which to deliberate together about which norm to adopt. Since the parties are jointly committed
to such an agreement, it is achieved, Hume says, as soon as there appears to them a salient focus for it.\textsuperscript{277}

This is important because rules may sometimes need reconsideration, when they stop bringing good consequences. Thus the source of obligation is not the rule, but the plural perspective from which we decide that this rule is necessary. The rule is good for a group of people, and this is why we adopt it, not because rules are important in and of themselves.

Postema thinks this is crucial, since it allows for critiquing the rules.

It may be true that without more or less determinate moral rules the plural perspective can yield no guidance for social interaction, but it is equally true that these rules can be intelligently and responsibly applied and followed only if they are understood and can be assessed by ordinary moral agents from the plural perspective.\textsuperscript{278}

3. Group Loyalty is the Natural Motive to Act According to the Convention

My goal so far has been to show that there is a partiality that underlies Hume’s conventions, and thus that the very impartiality of rules is not the source of impersonal obligations according to Hume. Thus this has so far been a negative argument, about what cannot be the source of impersonal obligations on Hume’s moral theory. I now turn to my positive account of Hume’s impersonal obligations. In order to argue that loyalty is the reflective impression that gives rise to idea of obligation in the case of impersonal obligations, I must show that loyalty can motivate a person to act according to the convention.

\textsuperscript{277} Postema, “Morality in the First Person Plural”, 39.
\textsuperscript{278} Ibid, 36.
Whether loyalty can motivate to act according to the convention connects to a discussion in the scholarship on Hume’s theory of justice, regarding the natural motive to act in accordance with the convention. As we saw in chapter 1, natural motives are ones that exist prior to any convention while artificial motives are ones that develop in response to the existence of a convention. Commentators on Hume agree that he thinks there is an artificial motive of respect for private property, which is developed in societies that have a convention of justice. According to the virtue theory reading of Hume’s account of justice, and according to my own interpretation, this is the approved motive whose approval gives rise to the virtue of justice. But all agree that Hume thinks we have artificial motives to act according to the conventions, so that we often return a loan out respect for private property, and we often keep our promises out of a motive of honesty. But Hume also asks, at the beginning of his chapters on justice, whether there is a natural motive to follow the convention, one that does not develop in response to the convention. It seems that Hume thinks there is no such natural motive, but since Hume isn’t clear about this philosophers have a field day arguing about it.

In this part I argue that the natural motive of justice is loyalty. I here follow Sayre-McCord in offering Hume a motive that he does not discuss to solve the problem he raises. First, I take a close look at two suggestions made by Annette Baier and Geoffrey Sayre-McCord, and argue that their views rightly indicate that trust and doing one’s share characterize the motive to follow the convention once it is up and running. But I argue that both of their suggestions miss out on the partiality that I argued is central to the adoption of the convention. This partiality, I contend, does not go away once the convention is in place, although we sometimes don’t notice

---

279 As discussed in chapter 1.
it. I argue that Humean loyalty is characterized by trust and responsibility, but better explains the natural motive of justice. Loyalty to a group motivates individuals to overcome self-interest or “parochial loyalties” and act according to the convention that serves the common interest of this group they are loyal to.

3.1 Common Interest and Trust

Annette Baier argues that what makes following convention possible is trust.\textsuperscript{280} Baier defines trust as the “acceptance of vulnerability to harm that others could inflict, but which we judge they will not in fact inflict.”\textsuperscript{281} We trust someone when we rely not only on their habits or usual actions, but on their good will, claims Baier.\textsuperscript{282} She uses Hume’s metaphor of two people rowing a boat together as her metaphor for trust. When two people row, they each row because they rely on the other to row, thereby each exposing his vulnerability to the other person. They only have reason to row because they know that they will get somewhere, given that the other is rowing as well.\textsuperscript{283} In the same way, Baier thinks that following the convention – repaying a loan, paying taxes – is motivated by trust. We accept monetary demands made on us, and thus accept that we are vulnerable to others and their demands. We only do this because we rely on others not to take undue advantage of this reliance, to respect our property by not stealing from us and returning the debts they owe us.

It is possible to construe trust as self-interested – it can be argued that one trusts because this is the best way to achieve something one wants. But Baier argues that trust is

\textsuperscript{280} Rorty reads Baier as suggesting that justice is a form of loyalty. See Rorty, “Justice as a Larger Loyalty”, 141.
\textsuperscript{281} Baier, \textit{Moral Prejudices}, 152.
\textsuperscript{282} Baier, \textit{Moral Prejudices}, 98-99.
\textsuperscript{283} Baier argues that Hume thinks promises are a “secured case of mutual trust”, \textit{Moral prejudices}, 111.
actually a motive that transcends the categories of egoistic and altruistic motives, assumed by philosophers.\textsuperscript{284} “Our actual motivation, in situations where trust comes into play, is not very helpfully seen as a mixture of egoistic and nonegoistic, unless we can be fairly sure which strands are egoistic, which altruistic. But many of our motives resist classification in these terms. Is a parent’s concern egoistic or nonegoistic?”\textsuperscript{285}

By exposing our vulnerabilities to others, Baier explains, the line is blurred between self-interest and other people’s interest.

We see why Baier thinks that trust, and the blurring of the line between egoism and altruism, is at the heart of Hume’s convention. Hume says that it is “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” (Treatise, 490). Hume says that convention depends on a common interest – an interest that is both mine and yours. The convention is possible because there is some shared interests, so that one convention can benefit all. Hume thinks the existence of common interest and its acknowledgement is what brings about the development of the convention. As we saw, the mutual acknowledgement allows for collaborated effort, “When this common sense of interest is mutually express’d, and is known to both, it produces a suitable resolution and behavior.” We can see how mutual acknowledgement of a common interest would give rise to trust.

But although Baier is right that trust is necessary for the institution of a convention, it is not clear that it is the motive which motivates people to follow the convention. Baier struggles to explain how trust is motivating. She suggests trust is motivating because of what it brings

\textsuperscript{284} Baier, Moral Prejudices, chapter 8. This fits with Hume’s account of the artificial virtue as depending on multiple motives. Although there may be one motive that we approve of, many other motives enable the cooperation.

\textsuperscript{285} Baier, Moral Prejudices, 156, 166-167.
about. She points out that trust exists in many of our relationships, and is motivating because it
brings about goods that cannot be achieved without it. But this does not really explain why
trust is a motive. It shows that trust is just a necessary condition for achieving other things that
we desire. Indeed trust is not a motive, but an enabler.

As an enabler, trust exists in all kinds of interactions. We trust a stranger to give us
directions, although our interests in this exchange are not common. When I ask for directions,
my guide’s motive is beneficence, while my own motive is to get somewhere, and it is mere
chance that the guide’s beneficence lines up with my needs. In these cases, trust does not rest on
the common interest that Hume points to as the basis of convention. Baier is wrong to claim that
in all instances of trust, there is a blurring of the line between egoistic and altruistic motives,
since the motive of the guide and the guided are clearly distinct. Hume doesn’t think conventions
grow out of such chance. We follow conventions because we have the same interest as other
people. “experience assures us still more, that the sense of interest has become common… And
‘tis only on the expectation of this, that our moderation and abstinence are founded.”

It is telling that Baier thinks trust is central to convention, but she also thinks that this
trust is prevalent in all human interaction. This means that we have endless relationships of trust,
and may constantly be developing new ones. There are endless forms of trust, because trust is not

---

286 Baier, Moral Prejudices, 100-101. “Since the things we typically do care about include things that we cannot
singlehandedly either create or sustain (our own life, health, reputation, our offspring and their well-being, as well as
intrinsically shared goods such as conversation, its written equivalent… and so on), we must allow many other
people to get into positions where they can, if they choose, injure what we care about…” She also suggests that we
trust because we find pleasure in knowing that others harbor good will towards us. Moral prejudices, 132.
287 For the distinction between reasons and enablers, see Jonathan Dancy, Ethics without Principles, Oxford
288 “We inhabit a climate of trust as we inhabit an atmosphere and notice it as we notice air, only when it becomes
scarce or polluted.” Baier, Moral prejudices, 98.
289 T 3.2.2.10, 490.
exclusionary: you can trust an endless number of people with different things. But as we saw, conventions are partial. The motive for such a partial interaction, which is prioritized over others, is loyalty. It is in these cases of loyalty, I contend, that our interests merge as Baier suggests, and the line between egoistic and non-egoistic motives-blurs.

3.2 Doing One’s Share in A Common Enterprise

Sayre-McCord argues that the natural motive of justice is doing one’s share. This is not a motive that Hume considers, but since Hume does not come up with a motive Sayre-McCord thinks suggesting a motive may be most helpful here. Doing’s one share is not about doing what is good for oneself, because it requires being aware that one’s actions are contributing to a larger enterprise. The motive, as I understand it, is a desire to be a part and contribute to bringing about a great good. Sayre-McCord thinks that once the convention is in place, and there are rules delineating responsibilities, the natural motive to do one’s share transforms into the approved motive of doing one’s fair share, since the convention delineates mutually beneficial rules.

Sayre-McCord is right that the motive to act according to the convention of justice must be a form of taking responsibility. The motive to act upon the convention is a motive to take part in bringing about a large good that cannot be achieved on one’s own. The motive to follow the convention is a motive to be part of a larger scheme. It is the motive to take personal responsibility, and not to expect that the good consequences of the convention will appear


without one’s active involvement. This motive is distinct from rule following, because it places the good to be achieved, rather than the rules, at the forefront. The motive of doing one’s share could appear without any rules – when people pitch in to make dinner together or contribute to a charity fund. But this same motive can then motivate people to follow the rules, once the rules are in place, if the rules are in place to bring about some good.

What I argue is missing from Sayre-McCord’s description is reference to the group with whom one shares the convention. The motive of doing one’s share only assumes that there is a good to be achieved, but doesn’t take into that this is a goal that is being achieved with other people, and that these people influence our motive. Sayre-McCord is focused on the common enterprise, rather than on those who it is common to. For human beings, who they share the enterprise with is quite important in choosing an enterprise and being motivated to keep contributing to it even when it is difficult to do so. When it comes to conventions, there is no common enterprise, independently of the group that shares it. The group precedes the convention, and thus the relation of a person to the group is going to influence the motive of contributing to the convention. Thus while I agree that there may be a personal trait of being motivated to do one’s share, the natural motive to follow the convention is more likely to be a motive that responds to the existence of a group that shares the convention, rather than just wishing to achieve the prosperity achieved through the convention with no reference to the group.

Sayre-McCord points to an important feature of the motivation of the artificial virtues, which display responsibility. But he does not account for the social context that causes such a motive of responsibility. This motive is not just a desire to participate in bringing about
some good. It is a response to the existence of a group one is part of and who is working together to achieve some good.

3.3 Loyalty to the Group that Shares the Convention

We started out with the question what is the natural motive to act according to a convention, once it is in place. I argued that Baier is right that keeping the convention depends on a merging of interests, since we are not motivated by a purely egoistic or altruistic motive, but I argued that what motivates can’t be trust because trust is an enabler rather than a motive. I further argued that Sayre-McCord’s suggestion of doing one’s share correctly points to responsibility in a joint effort, but is not sensitive to the question who are those we share the effort with. I this section I build off of Baier and Sayre-McCord’s suggestions, and argue that loyalty is the natural motive to act according to the convention because it can better account for both responsibility and trust. Loyalty motivates us to follow the convention because the convention serves a need of the group, and prevents the group from disintegrating into parochial loyalties.

Usually when we think of loyalty to a group, such as a society or a nation, we think of loyalty as the motive to defend this group from external competitors or enemies. Thus we think that loyalty motivates soldiers on the battlefield, or politicians in their international interactions. Hume does discuss external dangers as a motive to come together as a group, in his discussion of how governments are instituted. But Hume’s account of how society comes about emphasizes that the main danger to society is tearing apart from the inside. The danger is

292 T 3.2.7, 534-539.
that people will act on their self-interest or their narrow loyalties. Of course external dangers exist, and the person who is loyal to his society or nation may be called on to fight in a war. But the more common way in which a person is required to act out of loyalty is in placing limits on his own self-interest, thereby strengthening the common bond of the group rather than pulling the group apart. What allows for the group to stay together is the convention, which maintains the relationship between the group members and focuses it on aspects that they share and which they can collaborate on. The loyal person does not view these conventions as mere rules, but as the tools to ensure the existence of the group that she is part of. Thus when she is faced by a strong urge to serve her own self-interest, and to act against the convention, she is also motivated by loyalty to act according to the convention.

The partiality of loyalty fits with Hume’s view of the particularity of convention. A convention is not any mutually advantageous rule, but one that is found to be beneficial to a particular group through the joint reliance of this particular group of people. As we saw, Hume thinks that a convention is a rule that people adopt and act upon because they rely on others to act according to the rule, as well. But the expectations of others don’t always motivate us, particularly when they require that we sacrifice some of our interest or the interest of those we are close to. What makes the reliance of others into a motivation for one to act is one’s loyalty to those who have these expectations. As we saw, what motivates the loyal person is to fulfill the needs or strengthen the interests of the object of loyalty. What is the need that we respond to when we follow the convention? How does loyalty actually motivate to act according to the convention?
Sometimes when we act on behalf of a group, the need that we fulfill is clear. For example, a soldier who goes to war is clearly motivated by loyalty to his country or nation, and the need that he fulfills in protecting his country is clear. But the need that we respond to, when we follow conventions, is not so clear. This is precisely why Hume says that the benefit of society as a whole rarely motivates people to follow the convention. This is also why Hume needs to develop the category of the artificial virtues, which, as we saw in chapter 1, are not clearly beneficial in every instance. In the following example I will want to show that when we act according to conventions, the need that causes us to be motivated by loyalty is the danger of our group disintegrating.

Let us look at an example that Hume himself brings in the context of his discussion of justice. Hume brings the example of returning a loan to a miser, who will never use the money and who we feel no benevolence towards. In this example, I suggest, the motive is not loyalty to the particular individual, the miser.\textsuperscript{293} We have no loyalty to him as an individual, for we may not know him or have any past relationship with him. What motivates us to treat the miser according to the impartial rule of justice is our loyalty to the society or nation that we are a part of, which shares this convention of private property. We treat this miser not merely as an individual with a personal interaction with us, but rather as a member of a group that shares the convention. In this way, we are motivated not by the needs of the miser, but rather by loyalty to the group with whom we share the convention of justice.

Hume considers whether sympathy with the interest of society will motivate us in this instance, and he discards this option. This is because he thinks we can construe returning the

\textsuperscript{293} T 3.2.1.13, 482.
money to the miser as not beneficial to society in any way – the money is not returned to society, but to an individual. Hume is of course aware that constant fighting will cause suffering for society as a whole. But he thinks that sympathy with this suffering is far removed, and we don’t consider it often.294

But Hume misses the possibility that a society or group could have another interest, which is not a monetary one. Our society as a whole will not subsist if we act on self-interest or benevolence in such instances, for it would just fall apart into competing “parochial loyalties”. Thus there is an interest of the group to stay together which is harmed in an immediate, rather than a distant, way. It is precisely in this moment, where my own and my family’s interest conflicts with that of the miser, that the possibility of us living together in a larger society that works together towards prosperity is endangered. This is not an instance that is irrelevant or far removed from the purpose of the convention. It is precisely what the convention comes to ensure. In this instance the group is endangered by one’s own self-interest.

Thus the motive of loyalty in this instance is not caused by the needs of the miser or by the needs of society for prosperity. It is caused by the damage I will cause my society, by taking my own financial loss as more important than sharing in the cooperative effort of private property. I share with the other members of my city or country the joint effort of ensuring our joint prosperity. This joint effort led to the institution of certain conventions of behavior, such as returning loans. The benefits of such a convention can motivate in themselves. But this motivation seems weak in the particular situation I am in, when it is not a large sum of money and the miser won’t miss. But my joint effort with others will be hurt if I act against the rules we

294 T 3.2.1.11, 480-1.
together put into place. I will be betraying the group I am a part of, if I fail to act upon the convention, which articulates the ways in which we will look out for one another and work together to live peacefully.

It is true that trust characterizes my financial interactions with others. It is because people trust each other that they are willing to give loans. But trust is not a motive, it characterizes a loyal relationship. I respond to the trust of people with responsibility, because I view myself as sharing in an effort together with them. The motive to return the loans is loyalty to the group, and not their trust that I will do so. Similarly, it is true that I am being responsible by returning the loan. But I am not willing to be responsible, at great cost to myself, independently of who I am benefiting. When I follow the convention, often at cost to myself and to those who are close to me, I am motivated by the loyalty I have to a particular group of people, with who I share the effort of making our society better off.

3.4. A Comparative Note

If I am right, and loyalty is the motive to act on behalf of the conventions of the society that we are part of, then I am in disagreement with both egoists and sentimentalists. Hume’s view of how the self extends to include others leads us to a different view of how societies subsist. Egoists such as Hobbes and Mandeville thought we are motivated by self-interest to associate with a group of people, and it is self-interest that keeps people following the rules of society. Shaftesbury and Hutcheson criticized this view adamantly, insisting that human beings are benefitent and social in their nature, and that they find pleasure in the company of other human beings. For Shaftesbury and Hutcheson, loving one’s society is an expression of the
love of all humanity, and it is this universal beneficence that would motivate following the laws of justice.

Hume is quite critical of this view that human beings love other humans as such, and in his chapter on justice he flatly denies that there is such a sentiment. Thus he clearly does not think that this is the motive for creating societies and upholding them. Hume thinks that our love of other human beings is limited, and that we have strong preference for those we are close to. The love we have for family and friends cannot be extended to all of humanity, and so there is no hope that it will be. Hume agrees with Shaftesbury and Hutcheson that once society is founded, human beings discover that they enjoy the company of other human beings. But he doesn’t argue that this pleasure is sufficient to motivate keeping the laws of justice.

Hutcheson and Shaftesbury thought that the best motive is universal love. Shaftesbury has an extended discussion of what he calls “associationism”, being motivated more by the well-being of a sub-part of society rather than society as a whole. He thus seems to think that love which is not universal, love to a group, is dangerous to the well-being of society as a whole. Shaftesbury won’t be the first or last to express the concern that loyalty is non-moral or amoral because it is an exclusionary attitude, by which someone prefers those who she is loyal to over other people. This preferential treatment is often looked down upon by those philosophers who think that morality is fundamentally a universal love or equal treatment of all human beings. But if we share Hume’s view, that universal human love is not attainable and therefore not desirable, then we start thinking of loyalty in a different light. Loyalty looks good, when we realize that loyalty is what enables us to find the needs of others pressing as if they were our

---

295 T 3.2.112, 481.
own, so that we are not merely motivated in this way by our own egoistic needs. Loyalty is what enables humans to create families, friendships, teams and societies, in which we don’t have individuals looking out for their own egoistic benefits but acting to benefit others in a consistent way.

We are more likely to view a group as our own, than have beneficence towards every individual in the group. Associating someone else with the self means that we are motivated to look out for their needs consistently. This kind of association with the self can only happen if we are particular, and choose among people, teams, and societies. The very exclusion is what makes someone my family, my team, or my society; if there was no exclusion, there would be no association with the self and no responsibility towards one’s own. This would not lead to more altruistic beneficence, it would merely cut out the possibility of acting in these towards a large group of people.

4. The Psychology of Impersonal Obligations

So far I’ve argued that for Hume, the convention of justice arises from the experience of a particular group that faces a problem and seeks a solution so that they can live together. The rule is a tool, used to solve a problem faced by the group. I also argued that loyalty motivates us act according to a convention shared by our society, when our tendency to prefer our self-interest over the group gives rise to the danger of disintegrating the group. In the following I argue that it is this reflective impression of loyalty which give rises to impersonal obligation. Impersonal
obligations, like personal obligations, are explained by the psychology of loyalty. I then respond to the interpretation offered by Wiggin that all Humean obligations arise from impartial rules.

In chapter 3 I argued that obligation has its source in the impression of the motive of loyalty, which is most often observed in instances when there is a conflict of interest between self-interest and loyalty. I showed that in instances where one is not strongly motivated to act virtuously in one’s personal relationships, loyalty can act as a strong motive to act on behalf of those who are part of our extended self. This strong motive, which is distinct from moral sentiments, is the impression that is the source of the idea of obligation. Now I would like to argue that in the case of impersonal obligations, which Hume thinks arise from conventions, obligation to follow the convention still comes from loyalty, this time loyalty to the group that shares the convention. Although we judge conventions good, and we judge acting according to the convention (being honest or just) virtuous, the obligation to be honest or just does not come from this moral approval, nor from the impartiality of the rules. Rather obligation comes from loyalty to the group that shares the convention.

I have shown above how loyalty motivates to follow conventions, in moments where we have strong motives to act against the convention. What causes the motive of loyalty in such instances is the harm one poses to the group, by being motivated to act against the convention that keeps the group together. This harm is the coming apart of the bond that was created among the members of the convention, and it is posed by one’s own tendency to act according to one’s own self-interest or “parochial loyalties”. The motive of loyalty is to respond to this harm by taking up the well-being of one’s society or larger group that shares the convention.
This reflective impression of loyalty gives rise to an impersonal obligation. In the moment when we are motivated to avoid paying our debt, the consequences of all members of the convention respecting private property are far from our mind, Hume argues. But I suggest that it is specifically in moments when the good consequences of following the convention are far from our mind that we act out of obligation. While we are sometimes motivated by the good consequences of the convention, or at least some of us are some of the time, this kind of motivation does not explain obligation. It does not explain the experience of doing something “because I have to”. When we do act out of obligation we are motivated by loyalty to the group that shares the convention. What motivates us in these instances is not the rule of the convention, but the loyalty to those who share it. Such instances of loyalty to the group with whom we share a convention give rise to an idea of impersonal obligation, distinct from the virtue of honesty or acting on the interest of society.

On this theory of obligation, there aren’t two or three sources of obligation. Rather all obligations come from the reflective impression of loyalty. The difference between personal obligations and impersonal obligations is not in the impression that gives rise to the idea of obligation, but in the object that one is loyal to and thus the kind of actions that loyalty to it requires. In the case of personal obligations, loyalty is to a particular person, and thus the requirements of loyalty are to strengthen this object of loyalty. In the case of impersonal obligations, loyalty is to a group, whose interests a person cannot fulfill on one’s own. The convention spells out the ways in which a loyal person can promote the security and prosperity of the group. This security and prosperity can sometimes only be promoted by acting in particular ways towards members of the group.
It is worth noting that this view of obligation does not render acting out of obligation the moral way of acting. When a person is motivated by the good consequences of the just system, they are acting morally. When a person is motivated by the motive of honesty, which is approved from the general point of view, they are also acting morally. Thus a person does not need to act out of obligation, in order to be moral. But there is a distinct motive, on this view, which is “acting out of moral obligation”. What is distinct about fulfilling obligations is that it is caused by the harm people and groups that are central to the agent’s self. It assumes that human beings have the ability to extend themselves, and that obligations arise from this extension of the self to include others. It is not a motive that responds to a particular good that is placed before the agent’s mind right now – self-interested or beneficent. Rather it is caused by the impending harm or possible good that is impending on particular people or groups that are part of the agent’s self, and thus motivates the agent consistently.

The most serious challenge to this reading is one that views all obligations, even on a Humean moral theory, as arising from rules. The claim would be that when we are in a conflict about giving back the debt to a miser, the impression we have that motivates us to return the loan is that there is a rule that stands independently of our motives saying that “One must return loans”. What is suggested is that in a given moment, we do not want to follow the rule, and yet it stands there as a rule that dictates moral action, and this is what we experience in the instances when we act out of obligation. We can see why someone would think this. When we talk about obligation, we talk about what we do when we don’t want to. It seems easy to claim that we have a natural ability to take up a rule, which then will apply when we don’t want to act in the particular way the rule dictates. We can then say that when the rule dictates action, but we don’t
want to act accordingly, we are obligated. The rule has some pull on us, since we have the ability to apply rules to ourselves, and it is this application of rule, rather than a motive of following the rule, which gives rise to the impersonal obligation.

Although Darwall only claims that Hume holds this view in the case of justice, others have read Hume as holding that all obligations are generated by rules, according to Hume. David Wiggins has argued that we can find a categorical imperative in Hume, if we take seriously that Hume thinks human beings come up with standards of acting that they approve of from a general point of view. Although our own sentiments may not always match up to this standard, we are able to apply this impartial standard once we have it. As Wiggins reads Hume,

> At this point, namely the point marked by our assuming full mastery of “general language”, our imagination, our reason, our powers of analogy or of looking for similar treatment for similar cases and dissimilar treatment for dissimilar cases, the unending search for generality, have all cooperated with our benevolence, with our capacity to resonate the feelings of others (which Hume calls sympathy)… in a way that enables us to make the shared standard our own standard of assessment… there is no clear limit upon whose happiness or misery can impinge upon the judgments that issue from the public standard. For the standard is entirely general. 296

For Wiggins the fact that we have a shared standard defines obligations because it can be applied to any person, regardless of whether they share our judgment, by virtue of the fact that this person has sympathetic abilities which would enable them to share in the judgment if they took up the general point of view. On this view, for Hume “a duty is enshrined in the abstract rule or public standard for the assessment of characters and the actions that express the characters.” 297 In defending his Humean account of categorical imperatives, Wiggins argues that the fact that the

---

297 Ibid, 88.
standard is not *a priori*, but rather is the standard that we have come up with and share, does not mean that it can’t generate obligations. Its contingency on human nature perhaps worried Hume, but this does not mean it can’t be an explanation of obligation.

Wiggins’ attempt to show that obligations, on Hume’s moral theory, arise from applying moral rules is mistaken on two accounts. First, Hume purposefully distances himself from defining morality in terms of rules. This is clear in his account of moral judgment, which does not have its source in rules. Moral distinctions for Hume come from sentiments of moral approval and disapproval. It is true that in order to regulate these moral sentiments, and share our language with other people, we do develop a standard for what we approve or disapprove of. But the source of the distinction is not the rule, but the impressions that give rise to the rule. In the same way, I argue, obligation on a Humean picture cannot come from rules. We may systematize our impressions of obligations into rules, as we do with the impressions of moral approval or disapproval. But the source of moral obligation is a reflective impressions, not the rule itself.

Second, Wiggins’ account of how Humean obligations arise does not discuss Hume’s characterization of the problem of obligation. As we saw in chapter 3, Hume claims that we can explain moral obligation only in terms of a motivation that connects moral distinctions and the will. Wiggins’ account of obligation does not show what motivation connects between moral distinctions and the will to act, and thus it doesn’t offer a Humean explanation of obligation. The application of a moral standard may generate some motive to follow it, the application itself is not a motive that offers an explanation of obligation.

---

298 Ibid, 88-89.  
299 Ibid, 90-91.
I am not disagreeing that we often act according to rules, or that these help us figure out in what way we can be moral. My objection is that the application of a rule can give rise to an obligation, on its own. Can you be loyal to a rule? Perhaps this whole argument between obligations arising from rules and obligations arising from loyalty is a mistaken dichotomy, because we could have the motive of loyalty to a moral rule. On my theory of Humean loyalty, this is not possible. Loyalty is only possible when we have a particular entity – a person, a group, a tradition – that is associated with the self. What this association makes possible is that the success of this particular entity affects us as if it were our own, and its failure affects us as if it were our own. A rule cannot be associated with the self, because the rule is in essence something not particular, which does not succeed or fail, is not harmed or assisted. The loyal person responds to the harms and needs of the particular entity associated with herself. There is no way for a rule to have these needs or harms that the loyal person responds to.

It is precisely this partiality of loyalty that some think we try to avoid, when we speak about obligations. The Kantian influence in meta-ethics is that we try to explain obligations by our ability to act according to rules, which can be applied to all equally and can make demands on all equally. It is precisely this kind of universal obligation that does not exist, on my theory. Even the obligation to be just, on Hume’s and Rorty’s account, is an obligation that arises from being a member of a particular group with a particular history, and not from grasping some pre-existing rules that all human beings should follow. Here I agree with Rorty that “we cannot resolve conflicting loyalties by turning away from them all towards something categorically distinct from loyalty, namely the universal moral obligation to act justly.”

---

300 Rorty, “Justice as a Larger Loyalty”, 142.
is no obligation, except when the individual has a responsibility placed on her by her existing relationships, personal or impersonal.

5. Preliminary Thoughts on Normative Implications

In chapters 3 and 4 I argued that the reflective impression of loyalty gives rise to the idea of obligation, but I mostly avoided discussing the normative implications of this view. At least one of the outcomes of defining obligations is that we hold others accountable to fulfill them. Other theories of obligation try to explain how we can hold all human beings accountable by moral obligation. Underlying Immanuel Kant’s categorical imperative and other theories of moral obligation is a universalist assumption that all human beings are morally obligated by the same obligations.\textsuperscript{301} In this part I discuss one of the outcomes of my explanation of obligation, that we do not view every moral obligation as pertaining to all people.

I argued that the psychological source of obligation is not the ability to follow rules or the sentiment of moral approval, but rather loyalty. I’ve shown that loyalty can motivate us to act on the interest of an individual or a group that we are loyal to. This means that we can holds others morally accountable by obligations that arise from personal relationships, or from their membership in groups, institutions, and societies that they are part of. Obligations will include acts that strengthen the interests of the objects of loyalty from other people or groups that are endangering them. But more often than not, we will have obligation to act on behalf of our loyalties by overlooking our narrow interest in order to find the interest we share.

\textsuperscript{301} Wiggins account of Humean obligation, discussed above, has a similarly universal assumption.
My theory allows to hold people accountable for obligations, but it does not allow us to hold *all* people obligated by the *all the same* obligations.\(^{302}\) Rather the obligations an agent has arise from the associations that make up the agent’s self, and the motive of loyalty that is available to this agent in response to these associations. Thus we blame a person for cheating on his spouse, because his extended self includes his spouse and the motive of loyalty is available to him to remain faithful. We blame a person for not paying back a loan in a society that has conventions of justice, because her self is extended to include this society she lives in, yet she does not respond with loyalty to her society when she chooses to act out of self-interest. My claim is that the capacity for being obligated is common to all human beings, but that the particular obligations we have are relative to our existing loyalties.

My theory does not render all moral approval relative. I accept Hume’s view that we can morally evaluate people in all times and places. Many of his examples showing that we need a moral standard are empirical observations that show that do in fact morally evaluate individuals who lived long ago or enemies that we personally despise.\(^{303}\) The universal aspect of Hume’s moral theory arises from the assumption that we can, through sympathy, understand the pains and pleasures of other people, and that we can, through familiarity with human motives, ascribe motives to other people. Moreover, we can apply this moral standard to human beings in different social contexts, because human beings all share this sympathetic ability.

All this assumes a lot of psychological commonality among all human beings, and implies that at least with regard to the natural virtues, we can morally evaluate the characters of

\(^{302}\) My account focused on obligations to act in certain ways. It is possible that there are obligations to desist from acting in certain ways that do apply to all.

\(^{303}\) Hume discusses the need to transcend our self interest in order to maintain a standard of taste in “Of the Standard of Taste”, in *Essays Moral, Political and Literary*, 226-252.
all human beings universally. The same cannot be said for the moral approval and disapproval of
the artificial virtues. Hume emphasizes time and again the contingency of convention, arguing
that a society that is too small, or with abundant resources, would not need and therefore would
not have the convention of property that we are familiar with. Moreover, Hume argues that if a
society falls apart, because there is a severe scarcity of resources, then following the convention
will not be virtuous.304 Thus we cannot morally approve of a person for stealing or lying, if they
do not live in a society with conventions of private property or promise keeping.

Although there are categories of universal moral approval and categories of non-
universal moral approval on Hume’s theory, moral obligations are not universal on the Humean
account of obligation I have proposed. In the case of conventional obligations, Hume would
agree with me that moral obligations are contingent and not universal. Where Hume’s theory
clearly differs from universalist theories of obligation is that on his theory we a member of a
tribe is not morally obligated to return a cow taken from an enemy tribe. Where there is no
common interest that has given rise to a shared convention, a person is not obligated to act
according to justice. This is not only because a person cannot have conceptions of private
property, when they have no convention of justice, as Hume points out is the case. I contend that
on Hume’s view of obligation, my telling this person about the convention of justice that exists
in my society, and explaining the concepts that it gives rise, will still not give the tribe member
any obligation to return the cow to his enemy. Even if he morally approved of the convention
that exists in my society, and approved of my acting according to this convention, he would not
have a moral obligation to follow this convention. This is because at least with regard to the

304 T 3.2.2.20-21, 496-497.
artificial virtues, Hume thinks that our obligations are contingent on being members of societies with these conventions.

My own account of obligation goes further in relativizing obligations. Although I accept Hume’s universalist account of moral approval, I argued that moral approval that arises from sympathy, and obligation that arises from loyalty, play different roles in our psychology. Sympathy allows for universal moral appraisal, because it arises from our ability to sympathize with the sentiments of any human being to some degree. Loyalty is not something we can have to all humans, because it depends on partiality, and we cannot have a partial relationship with every person. On my account, while we morally appraise all human beings, in light of a universal standard of moral approval or disapproval, we only hold people obligated relative to those who are part of their extended self. Thus I could disapprove of another person not giving charity to a beggar, without judging that this person was morally obligated to do so.

Where Hume and I disagree is on the obligations that arise from the natural virtues. On Hume’s judgment internalist view of obligations, we are obligated to be beneficent because we approve of the motives that give rise to these virtues. Thus we can ascribe the obligation of beneficence universally. I objected to this view, since it entails too many obligations and does not explain how we prioritize among our moral approvals. On my own view, we can only be obligated to be beneficent when the beneficiary of the act is someone who is part of our extended self, because we morally judge our obligations by calibrating moral approval and loyalty. The disagreement here goes back to our disagreement regarding the source of.
As we saw, Hume is a judgment internalist about obligation. This means that in order for a person to be judged obliged to act in a certain way, there must be a motive in that person’s psychology that would explain how that person could be motivated to act on this obligation. Although our discussion of Hume’s judgment internalism was related to the way a person judges herself, the same can be said for judging the obligations of others. Just as there must be a motive internal to our own moral judgment that allows us to judge that we have an obligation, there must be an existing motive in another person, if we are to judge that the other person is obligated. The universal aspect of Hume’s theory is that we do assume that all human beings have motives to act virtuously, and we assume that all humans have the capacity, through sympathy, to evaluate themselves according to the standard of moral approval.

But I argued that if we take Hume’s judgment internalism seriously, obligations should not be ascribed to an agent based on the existence of moral sentiments alone. To ascribe to another person an obligation is to ascribe to them an ability to prioritize among motives in a particular way. Thus in order to say that a person must repay a loan when it is financially harmful for him to do so, it is not sufficient to point out that the person has, through sympathy, moral approval of returning the loan. Obligations do not arise from one motive, on its own, but rather from the ability to prioritize among motives. Thus we can only ascribe an obligation to someone, if that person has the ability to prioritize the approved motive or action. This prioritization does not arise from moral approval alone, but from the combination of moral approval and loyalty. Thus in order to ascribe to a person an obligation to return the loan, we must also point out that this person has an extended self that includes a society that shares a
convention of justice, and that he therefore can calibrate moral approval with the motive of loyalty available to him.

A question that I often get when I present my view of obligation is whether it can explain why Nazi S.S. officers were not obligated to kill innocent Jews during the Holocaust. Surely their loyalty to their country, or to the Nazi party, obligated these officers to commit these atrocities? I do not accept that my view entails this horrific conclusion in any way, or any other similar conclusion that ascribes obligations to commit immoral actions because of one’s loyalty. As I emphasized in the end of chapter 3, moral obligation is a calibration of moral approval and loyalty. On Hume’s account of moral approval there is no question that murder of innocent people is morally horrific. Given that moral obligations arise from the calibration of moral approval with loyalty, there is no way that SS officers could have had a moral obligation to kill innocent people. There cannot be a moral obligation to kill innocent people, even when this is demanded by one’s own country or party. Just as there is no moral obligation to do something evil, which will benefit one’s parent or spouse, there is also no obligation to do what is evil even if it is required by one’s country.

Another point to emphasize is that following the laws of any group or country is not necessarily loyal. Following the rules of one’s society is often loyalty, since rules makes it possible to live safely and prosperously with people we do not know. But this doesn’t mean that loyalty requires following all rules of one’s society. Some rules could be harmful to one’s society, and the loyal person’s end is the security and prosperity of his or her society, not following rules. The loyal person is motivated by the interest of their society or county to follow those conventions that are accepted and beneficial for their own society’s prosperity. The fact
that Nazi Germany passed murderous laws does not make them into conventions that are accepted over time because they are best suited to give Germans the possibility of prosperity and security. Clearly the murderous laws of the Nazi regime were meant to do the opposite for many German citizens. I would argue that a loyal German would have been motivated by his loyalty to Germany to oppose these laws that harmed the members of his society and turned the society he was proud of into a murderous one. Thus not only does my account of obligation not imply that murdering innocent people is ever morally obligatory, it also suggests that at times we are obligated to raise opposition to the rules of the groups or nations we are members of.

Some categories of actions that are considered obligations on other theories cannot be obligations on my own theory. One category that needs to be considered is that of obligation of self-improvement. On the theory of obligation I am offering there are no obligations to self, as such. It could be that I have an obligation to work on my pessimism, because it harms my husband and our marriage. It could be that I have an obligation to curb my avarice, because it harms my community and nation. But I do not have obligations to improve myself in ways that do not influence others. I may be motivated to work on my own vice of feeling inferior, because it causes me misery. But this moral disapproval of my inferiority complex, that motivates me to change, does not turn this goal into an obligation. While I may be motivated by moral disapproval or self-interest to follow through with this decision, when self-interest or moral disapproval are weak and cannot motivate me to act according to my resolution there is no additional motive of loyalty to strengthen their influence.

Another category of actions not obligatory on my theory are actions of benevolence to strangers. In this chapter we discussed obligations to act according to justice, to return loans or
refrain from trespassing. But many think that we have a more basic obligation to act with benevolence towards others. Ross refers to such an obligation, and Wiggins uses this example as the basis for his discussion of obligation. My theory does not allow that there is an obligation of kindness to all humans, as such. My theory allows that we have an obligation to give charity to a beggar on the street, but if we do have such an obligation it does not arise from the need of the beggar. There is no personal relationship that associates me with this particular beggar. Rather such an obligation arises if I live in a community or am part of a religious tradition that shares a rule of giving charity. This is not to say that it is not virtuous to give charity, even when one is not part of such a community. But the moral approval we have for such an act, on its own, does not make it obligatory. It becomes obligatory if we are part of a community that comes together to overcome poverty in this way.

What about obligations to those who are not part of our community then? The same should be said here. There is no obligation to feed someone starving on the other side of the world, just because they are in great danger. I think that if we hold that there is such an obligation, then we misunderstand the function of obligation to prioritize among endless demands on our time, strength and money. This use of the idea of obligation doesn’t point to another motive other than the approval we have of alleviating suffering. Thus it doesn’t explain how we would be motivated to act on behalf of the starving, in moments where the influence of moral disapproval is quite weak.

Although my view entails that there is no obligation to help those in danger just because they are in danger, this doesn’t mean that our obligations must be limited to those who are part of our community. Just as an individual is not limited to being motivated by self-interest
alone, a community is not limited to self-preservation. A community can dedicate itself to caring for others, who are not part of the community. When we understand that obligations arise from loyalty, we understand that obligations don’t arise from need but from ongoing relationships. If we are part of a community that works together to help those on the other side of the world, then we can have an obligation that comes from our membership in this community, to help others on the other side of the world.

**Conclusion**

In this chapter I argued that although there is a role for impartial rules in morality, they are not the psychological source of impersonal moral obligations. Impersonal obligations, like personal obligations, arise from the reflective impression of the motive of loyalty. In the case of impersonal obligations this loyalty is to a group, institution, society or tradition, which dictates that we treat its members or all human beings in a certain way. Having shown that loyalty can motivate us to act in accordance with conventions, just as it motivates us to respond to the needs of individuals we are closely associated with, I have shown that loyalty can offer a single source for both personal and impersonal obligations.
CONCLUSION

In this dissertation I presented a novel theory of loyalty, based in Hume’s psychology: loyalty is the motive to act on behalf of those who are part of one’s extended self, which is constituted by association of ideas of other individuals and groups with the self. I then argued that loyalty is the reflective impression that gives rise to the idea of obligation. I proposed that although moral sentiments explain our approval of actions, only calibrating moral sentiments with loyalty explains why we judge some actions not only good but obligatory.

This theory gives loyalty a central place in our moral psychology. It does not claim that loyalty is the only motive that forms our relationships, as is sometimes implied in the literature on loyalty. But it also does not view loyalty as just another virtue among many, as is sometimes thought moral theory. Loyalty gives rise to our obligations because it is the motive by which we prioritize some interests over others. This prioritization, when calibrated with moral approval, allows us to distinguish obligations among our moral approvals. This prioritization effected by loyalty functions as the strong motive to act morally.

This view has significant implications for moral epistemology, meta-ethics, and normative ethics. First, the view of obligation goes against a prevailing assumption that obligations arise from impartial rules. By arguing that moral distinction have their source in reflective impressions, Hume paved the way to a moral theory that does not view obligations as derived from rules. I developed this view, by arguing that we can point to a particular reflective impression that is the source of the idea of obligation. The position that obligations are actions that fulfill needs of those who are part of our extended self entails that we cannot know what our
moral obligations are by intuition or by universal rules. Judging that we have obligations requires knowing the facts about a particular person’s extended self, and this in turn depends on knowing facts about this person’s life history and social context. We can generalize such findings, since most people have obligations to their parents or friends who are part of their extended self, but the obligations come from loyalty and not from such generalization. Moreover, since being a loyal person is being motivated by the interests of a person or group who are part of the extended self, in order to know what our obligations are we must know what is in the interest of these individuals and groups. In particular, as Hume shows in his discussion of justice, what is in the interest of a group of people may not be possible to ascertain in any particular moment in time and may only be learnt through past experience. Promoting the interest of a group may require complying with conventions that can only be known to be beneficial in looking back on generations of such compliance.

Second, the view is a meta-ethical position that holds that moral normativity has its source in human psychology. In the terminology of this field, my view of obligation is a cognitivist, mind-dependent view. It is a cognitivist view of obligation, because there is a matter of fact that can be known about whether we are obligated to act in a certain way. It is not a realist view, as this terms is used in meta-ethics today, because the fact of obligation supervenes on other psychological facts regarding the mind of the person who is obligated, namely what is and what is not part of this person’s extended self. This is not a subjectivist view of obligation (as is sometimes mistakenly attributed to Hume), because obligations are not derived from the preferences of the individual. Obligations arise from the contours of the self, which are only
sometimes voluntarily drawn, but are always shaped over time and before the moment of obligation.

Third, this meta-ethical position has normative implications, since it implies that morally approved acts towards beneficiaries who are not part of the extended self cannot be obligations. This is true even if these morally approved acts are defined by some principle or rule that is held by the agent. This view is a unique, and may I say unfashionable, view in normative ethics. Although room is made in normative theories for agent-relative obligations such as familial duties, it is generally agreed that morality obligates us in obligations that are owed to persons as persons. My view is that although we morally approve of certain actions towards persons as persons, obligations do not fall out of such moral approval on its own. Rather obligations are only defined once we calibrate such moral approval with loyalty. This is because moral obligations do not come from our ability to evaluate, but rather from our ability to act morally.

There are four major areas that need to be developed to defend this view. First, some of the details of this theory need to be worked out. On the psychology of loyalty, I was not completely clear on whether the causes of loyalty are the needs of another, on their own, or whether the place of these needs within the cohesion of the self also plays a role in giving rise to this motivation. Second, the implications for normative ethics need to be spelled. I show that loyalty can explain personal obligations and impersonal obligations to act in certain ways, but I do not address other kinds of obligations. More specifically, I do not discuss obligations to refrain from certain action. These kinds of obligations do not fall into the format of prioritization,
which I argued is the function of obligation. I need to devote more thought to the status of such obligations on my theory.

Second, I only gave a sketch of how we weigh our different loyalties. I argued that when we decide how to act, we calibrate our moral approval of an action with our loyalty to the beneficiary of the action. But how do we decide among our loyalties, and weigh one obligation against another? I think we don’t run into these problems all the time, because we tend to extend our self to include individuals and groups who are loyal to each other. This is why tend to live lives in which our family members, friends, and co-workers are also part of a community, a tradition and a nation that we are all loyal to. But naturally this tendency is not always successful, nor does it do away with all conflicts among loyalties. Thus more needs to be done to study how we morally choose among our loyalties.

Third, I argued that moral judgment that calibrates moral approval and loyalty is cognitivist and truth-apt. But Simon Keller and Troy Jollimore, among others, have argued that loyalty is a prejudice that perverts our moral judgment because it blinds us to the moral failings of those we are loyal to and motivates us to prefer them over all others. I think the only way to respond to this challenge is to explain the role loyalty plays in helping us improve our judgments of moral approval. Further work needs to be done to understand what motivates us to correct our sentiments so that we come to have what we recognize to be moral approval. My work here was not focused on the epistemology of moral knowledge, but I believe that a Humean account will view socializing as an important component, rather than a prejudice, in our moral judgment.

Fourth, my theory of obligation rests on the assumption that Hume was right that we can advance in our moral understanding by studying human psychology. While I describe
Hume’s method, I do not discuss and evaluate alternatives to it. Further work needs to be done to distinguish the psychological approach from other empirical methods, such as the method utilized in phenomenology. I believe that the distinction will come from Hume’s division of perception into impressions and ideas. The claim would be that our moral concepts cannot be found on the phenomenological level of experience because our reflective impressions are not uniform. It is only when we try to find the common cause behind the impressions that we find an explanation for obligation.
BIBLIOGRAPHY


Abramson, Kate, “Happy to Unite, or Not?”, *Philosophy Compass*, 1 (3), 2006, 290-302.


Besser, Lorraine, “Hume’s Practical Conception of the Self”, unpublished manuscript.


Hazony, Yoram, “Induction in Newton’s Methodology”, manuscript.


