Knud Haakonssen has suggested in the opening chapter of the *Cambridge History of Eighteenth Century Philosophy* (Haakonssen 2006) that the history of the history of philosophy should be far more central to the history of philosophy than it currently is. Investigating and understanding how texts have been organized, edited, made available (or not), advocated for by their boosters, denigrated by their opponents with terms like “Hobbist” or “Schoolman”, and above all offered as curricula for the education of philosophers and historians of philosophy allows us to gain purchase on what past philosophy means for us. It also gives us greater understanding of why many writings by past philosophers that were thought to be philosophy in the past are now ignored or have even ceased to be considered to be philosophy. Beyond tracing influence and reception, the history of the history of philosophy involves reflecting on what roads are taken or not taken and why. Through the practice of the history of the history of philosophy we discover why we have the history we have and what (in part) this history is.

More concretely, I mean that “Locke”, when invoked by contemporary philosophers and historians of philosophy, is a grouping of passages and texts, a bundle of philosophical issues, literature on these philosophical issues by other philosophers and historians, and a smattering of biography and historical context. At the core of “John Locke” for philosophers is perhaps empiricism, personal identity, clusters of positions in the theory of perception and knowledge, the primary quality/secondary quality distinction, the nominal and real essences, etc., as well as particular problems and examples (i.e., Molyneux’s problem), and other philosophers such as Hume, Berkeley, Mackie, et. al. (understood again as bundles of texts and problems). And, depending on who is invoking Locke, contractualism, classical liberalism, property, Malebranche, Catherine Cockburn Trotter, or Thomas Reid, may be more or less peripheral. Further out are toleration, positions in religion, the philosophy of education, and issues and positions even more remote. Even when applied to passages, texts, and problems outside of these restricted contexts by very skilled and knowledgeable historians, these sorts of background assumptions often guide how “Locke” is taken up and understood. This will vary in different national and academic contexts, but that it happens is both beyond dispute. With a wide-ranging figure like Locke who is canonical in many areas, what is viewed as center and what as periphery reflects the priorities of the discipline as a whole.

---

1 Why the history of the history of philosophy is not central to some areas of the history of philosophy will be made clear.
In this paper I will first reflect quite generally on this process and then consider more particularly some of the special features of how this process occurs in ethics or moral philosophy. In the second part of the essay I will turn to a particular case -- a passage from Joseph Butler. I’m interested in this particular case because one worry about our practices as historians of philosophy is that due to our tacit reliance on these heuristic maps, and other issues I will discuss in the foregoing, it becomes very difficult to recognize when a philosophically influential reading of a passage has fueled the historical importance of a particular passage. This is because due to the reading the passage itself often takes on an historical importance that in turn makes it very difficult not to see the passage in the light of the reading. I will refer to passages of this sort as “emblematic” – i.e., emblematic of a philosophical issue or cluster of philosophical issues.

2. The History of Philosophy and Emblematic Passages

It might seem that I am targeting historians of philosophy who focus on the rational reconstruction of arguments. I am not. What I am describing is, to my mind, ubiquitous, generally irresistible (although locally resistable) and holds of all sorts of historians of philosophy in philosophy departments – myself included. Indeed I think that many historians of philosophy who are oriented towards casting arguments as clearly as possible in contemporary philosophical terms and then evaluating them tend to avoid two problems that many more “contextual” historians sometimes do not. They are open to the fact that problems and arguments might be more interesting than figures and schools and thus they avoid advocating for a particular historical figure (I will discuss this further below). And relatedly they often accept that arguments need not be defended at all costs, that they are often bad or false or even not philosophical by contemporary standards. This doesn’t mean that I’m advocating for this approach either. Rather, I want to underscore that what follows is not directed at one particular approach.

Most historians of philosophy are acutely and personally aware that the way they present history, and the history they are given to present, is indebted to the process of organization and selection I hinted at above. We produce syllabi and anthologies that need to make choices of inclusion and exclusion. And we are critically aware of the process, to a degree. For example many or even most of the historians of philosophy I know who teach the standard “British Empiricism”, “Continental Rationalism”, and “History of Modern” courses offered in many philosophy departments try to modify the structure of their course to include figures other than Descartes, Hume, Locke, and Leibniz as well as a broader range of topics and problems – some non-white non-male authors, some moral or political philosophy, a few issues connected with theodicy, etc. Historians who work on ancient philosophy, Medieval philosophy, and on the history of non-Western philosophy tend to be more aware of it than historians of modern philosophy often are because commentary traditions, shifting ways of interpreting texts, and exclusion and restriction are very
much on display in their practices as historians.

But even with awareness, basic frameworks are still set by curricular requirements imposed by the institutions we teach in and work in, or self-imposed by our sense of what is the core knowledge that students need to acquire. This process also determined and determines how canonical passages, sentences, phrases and bits of text, are understood in relation to these histories and figures, and how the emblematic importance of these passages – as bearers of weighty problems or decisive arguments -- provides warrant for the reader not reflecting on their history. I will discuss this at greater length in a moment.

First though, I am describing the way in which contemporary trends in the disciplines clearly have and have had a great effect on what historians choose to write about and how they choose to present what they choose to write about. The reasons for this are not mysterious. Most historians of philosophy are educated in contemporary philosophy in addition to their specialties and this informs – sometimes decisively -- how they think about their discipline. For example what draws historians of philosophy to study their discipline is often closely connected with their interests qua contemporary philosophy. If they are trained in history within a philosophy department it is often tailored to and in service of these philosophical interests. And, generally, they find the subject matter interesting qua philosophers, and what they find interesting is often a consequence of their orientations towards contemporary philosophy or contemporary presentations of figures and problems in the history of philosophy.

For example, if my interest in the history of ethics is in tandem with an interest in contemporary virtue ethics I will be motivated to investigate and present the past on the basis of a contemporary philosophical positions that I use as to orient myself towards, restrict, cull, and organize and make sense of past texts. I may believe that the philosophical position I hold, or more precisely the cluster of beliefs I have about virtues and happiness and so forth, was held by Aristotle, although more likely I am wholly aware that Aristotle’s natural slave argument or his discussions of women and so forth are not positions I maintain. But it is still likely that I hold 1) that there is a core view called Aristotelian virtue ethics that is compatible with my beliefs despite some notable differences, 2) this is the “philosophical” core of the position held by Aristotle, and 3) how I exclude elements from the core or even the passages that jump out at me as important are likely a consequence of my beliefs qua contemporary philosopher.3

---

2 I’m using contemporary philosophy in a very vague sense here to include many kinds of current philosophy, issues that seem currently important or pressing, etc. 3 As to 3) many historians of philosophy have had the experience of reading a very familiar passage and noticing a sentence or passage for the first time, despite having read (and perhaps taught) the passage many, many times. This experience points both the steady operation of 3) and the fact that all of the phenomena I am
I do not wish to deny that moral philosophers have been influenced by Aristotle’s actual views or by reading Aristotle. Rather there is a culling of acceptable views or interesting contents in connection to preferred elements of contemporary positions – in this case “virtue ethics”: a responses by some mid-twentieth century (and later) moral philosophers to what they saw as limitations and defects in the range of contemporary positions in moral philosophy – which often goes hand in hand with advocacy for the position. And I wish to suggest that this is happening in tandem with an accretive process whereby the choices that come to mind – emblematic passages, authors, and problems – are further restricted by the repeated iteration of this process -- the accretion of the stock of the history of philosophy according to the needs and interests of historians of philosophy are often organized – knowingly or not -- around the views and interests of well-recognized and/or powerful philosophers. Because historians of philosophy tend, and have tended, to be interested in philosophy as understood by their contemporaries, because of the esteem they and others have for their most brilliant philosophical contemporaries, and because of the importance these contemporaries have in orienting philosophical interests the ways in which philosophers who have great stature in the profession have represented their history has a far more pervasive and long-lasting influence on this process than how historians who are not well-known or powerful philosophers have presented it. We just take the end result of this process as “what’s philosophically interesting”. 

There is also, in parallel, an advocacy for the positions – either the importance of the philosophers and problems, or for these positions as true and philosophical. Consider how rare it is to meet a historian of philosophy who does not believe that they hold a variant of the views expressed by the philosophers they work on. Many identify deeply with the philosophers they work on. Again it is obvious why this is the case. Historians of philosophy are drawn to the views minimally because they think they are philosophically interesting, but also often because they hold them to be true. Furthermore they need to demonstrate to their colleagues who have no historical interests, or sideline historical interests, as well as other historians who work on different areas that what they have invested so much time and effort into is worthwhile. I’m not suggesting bad faith, far from it. Rather, I am suggesting that there is a tendency to defend one’s historical choices philosophically and this leads to advocacy for an author or a corpus as history and the most successful language of advocacy is the language that one’s fellow philosophers recognize.

What I am describing is often not so active. There is a stock of choices, the stock is very interesting, and there’s often little motivation to go beyond the stock and ask about the content and structures one’s own biases (what’s interesting) and the motivation (why). The literature on canonical topics, Hume on causation for describing are for the most part. I’m not suggesting anything like absolute closure. Also as noted above I think many people who work in Ancient philosophy are particularly aware of the history of the history of philosophy.
example, is often sophisticated and deep, and deeply historically intertwined with issues of contemporary interest. I would say that topics like this have been selected for reasons of their intrinsic philosophical interest, but in many cases it is difficult to separate their intrinsic philosophical interest from the history that has followed from them, and difficult to separate all of this from what philosophy is (they are central to what philosophy is) and how it changes as such. Indeed it is a kind of history of philosophy as contemporary philosophy. One does have to narrow one’s interests, obviously, but it seems how and why one narrows one’s interests and how much of it involves extra-disciplinary and arbitrary factors are worth some reflection. The history of the history of philosophy provides one axis for reflecting on this process.

There are other influencing factors that reinforce what I am describing of which I will only mention a few. The avenues of publication taken seriously by philosophers in general and the comparative respect afforded publications in top-tier non-historically oriented journals vs. historically oriented journals, both militate against certain types of historical work. Top-tier generalist journals tend to be tipped, understandably, towards history of philosophy that may be of interest to their more general (and thus non-historical specialist) readership and so is focused on the sorts of problems and topoi I have mentioned. A related effect is the consequence of the comparative esteem afforded different areas of contemporary philosophy: aesthetics vs. metaphysics for example. Furthermore factors like professional esteem which has a direct effect on jobs and appointments, the aforementioned esteem in the eyes of one’s non-historian colleagues which has an effect on salary and quality or life, and concern for a doctoral students’ future prospects if they do not work within a recognizably philosophical area and exhibit some of their skills in terms of contemporary philosophy. The result is a kind of self-selection and also self-censorship. This is not to suggest that views do not change, they obviously do. But the changes are often driven by contemporary philosophical views.

Some of these are endemic to academia as a profession and some are special problems in philosophy. But there are two special problems in the history of ethics that I would point to. First, moral philosophers often have a different attitude towards the positions they adopt and their arguments than philosophers in other areas. If I write a paper on a view in metaphysics, for example I argue in the paper that anything can form a composite with any other thing and this composite gives rise to a new thing, there may be a great deal at stake intellectually, but less personally. If you criticize my view I may be thin-skinned and I may feel you’ve attacked my abilities as a philosopher. But the views of moral philosophers are often more intimately connected with how they understand themselves not as philosophers but as persons. They reflect what they understand to be a worthwhile moral agent, character, or action and why it is at stake. To undermine moral views is sometimes to suggest that the holder of the view is morally questionable or a less than a decent person – for example repugnance objections.

Why I think this is the case is connected with a further restriction. Contemporary
moral philosophy is often subject to a restriction on worthwhile philosophical positions by the intuitions and beliefs of a very narrow group – academics (mostly white, developed-world, male) – about morals and moral subjects worth pursuing independent of philosophical reflection. This is not a recent phenomenon, but I would like to suggest that one of the consequences is that what is considered to be an ethical position or a figure worth investigating in the history of moral philosophy may be both more restricted and more contingent than what is considered to be a position worth investigating in metaphysics or mind.

I suggest that as a consequence this makes what is tendered as plausibly “moral philosophy” even more restricted in a given era than in other “core” areas of philosophy. Think for example of times when moral philosophy was thought of as the same as natural religion, or between odd it has been in some periods to think of ethics as part of political philosophy (and not in others – Aristotle, Marx, Rawls, etc.). It also often makes past examples of non-canonical “moral philosophy” seem much stranger to readers from a different era than discussions of metaphysics or mind. When settled intuitions about moral practices and the kinds of moral philosophy that take up and analyze these moral practices change they just seem like what they are describing is not moral philosophy at all and when philosophical fashions change whole areas are just sent of the pitch. For example think about the comparative neglect of seventeenth century moral philosophy over Hume, Kant, and Bentham.

These special restrictions in the history of moral philosophy when taken together – the proximity of what we identify as our moral beliefs to what we are and the force of contingent intuitions in guiding the philosophical analysis – tends to make the history of the history of philosophy even less wanted here than elsewhere. Since the history of philosophy is mixed with contemporary philosophy, for the reasons I have suggested, it is somewhat desirable by philosophers and historians of philosophy. But the history of the history of philosophy is a kind of meta-endeavor that is not clearly philosophical and so less desirable.4

That said a few authors and a few passages do persist, some for a very long time and some less so. The passages are what I have referred to as emblematic passages and essential ingredients in the construction of figures (like “Locke”), problems, and even movements. Because of shifts in what counts as philosophy and what is philosophically interesting, those that persist have a very special role in unifying the discipline and its self-understanding. One can think of a number of famous and coarse-grained examples – Aristotle’s natural slave argument and Plato’s noble “lie” have been emblematic for aeons but what they are emblematic of and how they are emblematic has changed in response to contingently connected historical changes –

---

4 A fair response might be, this all sounds Nietzschean/Foucaultian, which sounds pretty philosophical! The methodological considerations may be in line with the arguments of some philosophers, but actually engaging in the practice of the history of philosophy seems quite something else.
the use of Aristotle’s argument in the antebellum South and comparisons made in the twentieth century between totalitarianism and Plato’s ideal city. These are extremely famous emblematic passages that have had an effect beyond philosophical schools. Some less famous passages have only local interest to fragments of academia. The roles of passages change, often very subtly, as contemporary philosophical concerns change and how what problems the philosophers are discussing and what the philosophers are discussing change accordingly. And as the range of what is philosophy changes – i.e. natural religion or natural philosophy no longer are no longer philosophy so they are more rarely consulted by historians of philosophy in understanding emblematic passages -- the role of these passages changes.

I’ve painted this picture in broad strokes, but the main point is intellectual choices happen in a world of institutions and biases for us much as they did for our predecessors, and the choices we make both reflect our honest intellectual interests and a wide range of biases and restrictions. Indeed although these are conceptually distinct they are not separable in practice, they are mutually interconnected and reinforcing. The history of the history of philosophy makes this apparent in ways that the history of philosophy does not. Our own views about the history of philosophy are formed very much like those of past philosophers and historians of philosophy at points during periods of the dominance of academic philosophy, within institutions and responding to a wide variety of causal factors. One way to respond to this problem is by focusing narrowly on the history of the role of particular emblematic passages in the history of moral philosophy and thinking about how they’ve come to have the functions that they have and why we read them the way we do. This is what I will undertake a brief example of this in the next section. Be warned this is not a presentation of a smoking gun. To paraphrase Sterne, this is a too-quick account of how an emblematic passage has swum down the gutter of time.

3. An Example: Joseph Butler’s fifth objection

The appendix to this essay is the final objection (hence “the fifth objection”) from Joseph Butler’s “Of the Nature of Virtue” (ONV) (commonly known as the “Dissertation on Virtue”), and the discussion that follows it. “Of the Nature of Virtue” was one of two dissertations appended to the Analogy of Religion. “Of Personal Identity”, the other dissertation, is also emblematic. They are today normally discussed and taught in distinct philosophical contexts – criticism of Locke and (as we shall see) deontic criticisms of utilitarianism – but it is notable that they were both originally part of a work arguing for the probable evidence of natural and revealed religion based on inductive arguments drawn from analogies with natural processes. When I say “part of” I mean it literally. Although appended in the printed editions the dissertations were originally parts of the chapters “Of a Future Life” (OPI) and “Of the Moral Government of God” (ONV) respectively. Butler removed
them because they detracted from the main subject of the *Analogy of Religion* insofar as they are not really support for analogical arguments (Butler 1736: 438).

ONV (and within it the fifth objection) is brief (brevity is important for emblematicity) and one of a number of emblematic quotes, stretches, or passages in Butler’s work.\(^5\) It has led, in conjunction with a few passages on conscience in the first few of Butler’s *Sermons* and a few passages on forgiveness, to Butler being taken as a kind of British proto-Kant (Schneewind 1998: 522), a Kant without the full-fledged rational machinery but with an inkling of the internal ought. The first quotation in Sidgwick’s *Methods of Ethics* is from this passage,\(^6\) and Sidgwick viewed Butler as both a central inspiration to his way of thinking about moral philosophy in re the “dualism of practical reason”\(^7\) and as having offered powerful criticisms of utilitarianism that he felt had to be responded to (Crisp 2015: 2). In addition Sidgwick’s methodology has a strongly Butlerian flavor. It is notable that when Anscombe attacked the Sidgwickian tradition in ethics in “Modern Moral Philosophy” the first philosopher she tossed into the flames was “stupid” Butler (Anscombe 1958: 2).

A notable example of the fifth objection being taken as emblematic for my purposes is John Rawls’ mention of the passage in his *Lectures on the History of Ethics*. The *Lectures* were posthumously published and edited by one of his most influential students, Barbara Hermann. This is a long-standing pattern in the history of philosophy, that lectures on the history of philosophy by philosophers viewed as great, as central to the canon or as making a claim to being central to the canon – whether the sayings, lectures and essays of a number of ancient philosophers, Benjamin Whichcote’s *Sermons*, or Hegel’s *Lectures on the History of Philosophy* – are posthumously edited and preserved by the students and disciples who were influenced by them and believe that others ought to have access to them.

In Rawls’ *Lectures*, Butler is briefly discussed twice in the context of Hume. Here’s the more extensive of his discussions:

---

\(^5\) The most emblematic quote in Butler is no doubt “everything is what it is and not another thing” (Butler 1729, §39) although it has lost any connection to its context since being used by Moore as the epigraph to *Principia Ethica*. That Moore choose this epigraph says a lot about Butler’s foundational place in the British moral philosophical tradition.

\(^6\) Or more precisely the paragraph following the fifth objection where Butler reflects on the fifth objection. “The happiness of the world is the concern of him, who is the Lord and the Proprietor of it: nor do we know what we are about, when we endeavour to promote the good of mankind in any ways, but those which he has directed; that is indeed in all ways, not contrary to veracity and justice.” Sidgwick 1907: 3, quoting ONV §9.

\(^7\) That practical reason both justifies egoism and utilitarianism, and cannot adjudicate between them. See Schultz 2015 for a sense of the expansive literature on the topic.
To his *Analogy of Religion* (1736), Butler attached a short appendix titled “On the Nature of Virtue.” In this he argued, among other things, that many of our conscientious moral judgments do not seem to be guided by the principle of the greatest balance of happiness. Rather, our conscience, which Butler views as authoritative and regulative of our nature, is such as to “condemn falsehood, unprovoked violence, and to approve of benevolence to some preferably to others, abstracted from all consideration, which conduct is likeliest to produce an on overbalance of happiness or misery” (section 8).

In such a manner Butler believes God has framed our conscience and we are to act accordingly. Butler entertains the purely hypothetical possibility that God might follow the principle of greatest happiness; but even if so, that does not change the fact that our conscience as framed by God is to be our guide. We are to follow conscience.

“Section 8” is the aforementioned fifth objection. Rawls likely read the passage either in the widely read anthologies of British moral philosophy edited by D. D. Raphael (Raphael 1969) or by L. A. Selby-Bigge (Selby-Bigge 1897). Raphael followed Selby-Bigge in his selections from Butler. Schneewind also seemed to follow them in his anthology, but with the exception of the inclusion of the entirety of ONV as opposed to just the passage in the Appendix.

Selby-Bigge, today best known as the editor of a standard edition of Hume, studied at Oxford and gained “firsts in classical moderations (1881) and in *literae humaniores* (1883)” (N. D. Daglish 2004). Butler’s *Analogy* and *Sermons* were both central in (respectively) the theology and moral philosophy curriculum at Oxford in

---

8 Rawls’ quote from Butler is cited in the printed text of the *Lectures* from Jerome Schneewind’s collection *Moral Philosophy, From Montaigne to Kant* (Schneewind 1990) which presents the last of the five observations as an extract from ONV along with the “Preface”, the first three of Butler’s *Sermons* (“Upon Human Nature”) and selections from the sermons “Upon Compassion” and “Love of One’s Neighbor”. It is likely that Rawls was not citing the passage from Schneewind but rather that Hermann had located the citation in Schneewind in order to provide a readily available text for readers of Rawls’ *Lectures* since Schneewind’s anthology first appeared in 1990.

9 ONV was long commonly read appended to an edition of the *Sermons* and far less likely to be read in their original context appended to the *Analogy of Religion*. The *Analogy of Religion* was scarcer in philosophy departments in the mid-twentieth century US than it had been in the previous century. E. C. Mossner’s edition of the *Analogy* even left the ONV and OPI out. Presumably this was because ONV had migrated over the course of the nineteenth century to the *Sermons* (although an egregious editorial lapse all the same given that ONV was initially part of the *Analogy*, and Butler stated this in the *Analogy* itself).
the later eighteenth and throughout the nineteenth century. Indeed Butler’s *Sermons* was the sole modern moral philosophy text in the Greats curriculum after the reform of 1853 (Walsh 2000: 313). It appears that the long-standing role of Butler’s moral philosophy at Oxford was as a Christian moral counterpart to classical authors, and Selby-Bigge’s selection reflects this in spite of the fact that he read him as a paradigmatic anti-metaphysical moral psychologist.

Selby-Bigge grouped Butler with Shaftesbury and Hutcheson as the “three principle texts of the Sentimental school” (Selby-Bigge 1897, I:vi). Indeed Butler was for Selby-Bigge “the most typical of the British Moralists” (“Introduction” §6) and with him “the sentimental School really reaches its climax” (“Introduction” §20).10 The selections from the sermons “Upon Compassion” and “Love of One’s Neighbor” justify presenting Butler as a sentimentalist. Samuel Clarke, whom Butler greatly admired, is with the rationalists in volume II. This division was motivated by the fact that Selby-Bigge was a great admirer of Hume and structured his anthology to reflect Hume’s view of his predecessors. Butler does discuss moral sentiments at great length.

Selby-Bigge’s anthology set a pattern for how the British Moralists are read in general, and Butler in particular, insofar as non-specialist philosophers if they know Butler first-hand are likely know him via this anthology or anthologies patterned on his anthology. But the Butler texts that Selby-Bigge selected reflected a prior selection by William Whewell, today primarily known as Mill’s opponent on the question of induction in science but a dominant figure in nineteenth century moral philosophy.11

Whewell was one of the most influential figures at Cambridge in the mid-eighteenth century, which was with Oxford the most influential academic institution in the Anglophone world. Whewell was deeply disturbed by the fact that William Paley’s *Principles of Moral Philosophy* had become the sole text of moral philosophy required in the Cambridge curriculum. Cambridge was dominated by Newtonians, and centered on what we would call mathematics and natural science in distinction from philosophy. As Sidgwick notes, this was philosophy in late eighteenth-century Cambridge. In 1772 Jebb listed the four branches of philosophy as “Mechanics, Hydrostatics, Apparent Astronomy and Optics” (Sidgwick 1876). It only offered one Tripos (the honors bachelors degree with set subjects, readings, and exams) up to the establishment of the Moral Science Tripos in the mid-nineteenth century.12

---

10 These kinds of typologies are highly flexible. Notably Adam Smith (who follows Butler in the Selby-Bigge anthology) placed Shaftesbury with Clarke and Wollaston as a rationalist in his own survey of moral philosophy in TMS VII (Smith 1776: VII.i.1).

11 Beyond the choice of Butler selections, Selby-Bigge – the promoter of Hume – had wholly different philosophical commitments from Whewell.

12 For a detailed discussion see Gibbens 2001.
Sidgwick's utilitarian philosophy was an ideal moral philosophical text to those in a curriculum structured by the likes of Jebb.

But Whewell was wholly opposed to Paley and as he ascended at Cambridge he sought to dislodge Paley's dominance with his own moral philosophy. Butler, the untouchably credentialed Anglican Bishop was already very much central to the Oxford curriculum. He provided the ideal vehicle for Whewell. In 1837 Whewell gave four lectures, the express purpose of which was to substitute Paley for Butler (Whewell, ed.1865: iv.). The lectures were extremely successful. Whewell succeeded in getting Butler's moral philosophy into the Cambridge curriculum and in upending the dominance of Paley at Cambridge.

What was the general view that Whewell wanted to substitute for Paley? In his Lectures on the History of Moral Philosophy Whewell distinguished moral philosophy into two antithetical positions – Dependent Morality and Independent. Dependent Moralists “assert it to be the law of human action to aim at some external object” and Independent Moralists “would regulate human action by an internal principle or relation, as Conscience, or Duty” (Whewell 1862: 1-2). Whewell proceeded to describe two sides in a historical moral conflict with the Epicureans, Hobbes, Bentham, and Paley on the one side and Plato, Butler, and himself on the other.

It’s a precursor of the internalism/externalism distinction combined with normative anti-utilitarian commitments, but Whewell used the distinction as a principle to edit what Butler was essential to be read. He suggested an arrangement of the Sermons, which were the most important and their order, which forms the basis for the various anthologies described above. He also published his own redacted Butler that presented Butler as primarily an independent theorist of conscience (Whewell, ed. 1865). Butler’s work was taught at Cambridge before Whewell, but Whewell managed to replace Paley with Butler in the curriculum both prior to and then with the establishment of the Moral Sciences Tripos.

Sidgwick followed Whewell as a dominant figure in moral philosophy at Cambridge and of course in ethics to the present day. He was not an admirer of Whewell, and sought further reforms and changes in the system. Butler became emblematic of a utilitarian/anti-utilitarian conflict, and the fifth objection brought the issues to focus. Sidgwick was also a great exponent of Butler as previously mentioned – Butler was ubiquitous in this curriculum – so he was trying to both appropriate him as a paradigm formal philosopher and someone who deepened utilitarianism. Sidgwick argued that Butler was preferable to contemporary intuitionists, while at the same time attacking what he shared with Whewell (on the conflict in general see Schneewind 1977; Snyder 2006; Cremaschi 2008). This further cemented his central influence at Cambridge, as well as the central normative (utilitarianism vs.

13 In referring to it this way I am talking the language of post-Sidgwickian moral philosophy (and far later). Attacking confusions in Whewell was of course a central motivation in Sidgwick's epochal and epochally careful work.
deontic intuitionism) and metaethical (externalism vs. internalism) issues that dominated through Moore and onward. It also set how philosophers read Butler to the present day insofar as the controversies which the passage emblematized persist.

Now to move to Butler, who is he attacking? He is attacking those theorists who take benevolence to be the whole of virtue. The obvious candidate would be Francis Hutcheson (Schneewind 1998: 351-2) since he indeed took benevolence to be the whole of virtue, and his major philosophical works where he argued for this had appeared long before 1736.14 A second candidate, and I think even more likely candidate, would be John Gay, whose “Preliminary Dissertation” appeared attached to Edmund Law’s translation of William King’s De Origine Mali in 1731. Butler would very likely have been aware of this work before 1736.15 In the “Preliminary Dissertation” Gay argues for a theological utilitarianism that argued that happiness is the sole moral value to be promoted.

It seems obvious to readers that Hutcheson is the target because: 1) he stresses benevolence and 2) he comes directly before Butler in volume one of Selby-Bigge as a fellow Sentimentalist. Gay is in an appendix of volume two. I think Gay more likely because Hutcheson argues that the partialities we have for our family members is a source of benevolence warranted by the moral sense and that there are outward expanding spheres of benevolence is not problematic (Hutcheson 1723, passim). Hutcheson held like Butler that we had both a local guide to morality, our moral sense, and a more general one, our knowledge that God seeks for the happiness of the system as a whole. He thought surprisingly little about how the two potentially conflict. Butler’s opponent, though, is criticizing any particular partialities for the near and dear as in conflict with the one moral value: happiness. Gay criticized Hutcheson on just this point, that the moral sense which for Hutcheson provided a proximate guide to benevolence towards the near and dear was a mere occult quality, and all actions were resolvable into general reasons (Gay 1731: xiii-xiv).

But whether Gay or Hutcheson, I wish to further suggest that the passage has Samuel Clarke in the background as well. Clarke was of course a major moral philosopher and theologian, and a major inspiration to Butler, although the paradigmatic Rationalist in volume two of Selby-Bigge. Butler had corresponded with Clarke when he was still at grammar school16, and although he was critical of Clarke’s proof of the existence of God, he shared many of his substantive insights. The Sermons opens with a famous discussion of two methods in ethics that are complimentary, the former from abstract relations of reason, i.e. Clarke, and the latter empirically discoverable aspects of our moral frames or matters of fact, i.e. Butler (Butler 1729, vi-vii). As the Analogy of Religion is a work of philosophical

---

14 Irwin also suggests Cumberland as a background (2010: 521).
15 William King was perhaps Butler’s most illustrious predecessor as Bishop of Derry.
16 See Garrett 2013.
theology it would be surprising if Clarke’s writings are not in the background. The fifth observation refers the reader back to *Analogy of Religion* I.vi that in turn refers the reader to ONV, and as we will see this chapter from the *Analogy* discusses Clarke.17

How does Butler’s criticism go? Roughly

1. Benevolence (and the lack of benevolence) cannot be the whole of virtue.
2. If benevolence were the whole of virtue then we would only approve and disapprove of degrees of benevolence.
3. Imagine two competitors – Peter and Paul (AG – my names) – for something of equal advantage to both of them.
4. If a stranger aided Paul, that would on this account be virtue insofar as the stranger would be benevolent in making Peter more happy (considered independent of distant consequences).
5. Now suppose a stranger robs Peter to benefit Paul “who he thought” would benefit enough more than Peter that the benefit would overmatch Peter’s pains.
6. Furthermore suppose there are no further bad consequences.
7. The action would still be vicious insofar since we are constituted in such a way as to condemn these actions as vicious as well as to be partial in our benevolence independent of “which conduct is likeliest to produce an overbalance of happiness or misery.”
8. “And therefore, were the author of nature to propose nothing to himself as an end but the production of happiness, were his moral character merely that of benevolence; yet ours is not so.” ONV §8
9. “Upon that supposition, indeed, the only reason of his giving us the above-mentioned approbation of benevolence to some persons rather than others, and disapprobation of falsehood, unprovoked violence, injustice, must be vice in us, and benevolence to some preferably to others, virtue, from all consideration of the overbalance of evil or good which they may appear likely to produce.” Ibid.

In their interpretations of Butler, Whewell and Sidgwick (1906: 86) focused on the conflict between “intuitive” and utilitarian views. The fifth objection is almost always read as highlighting this conflict as a problem for utilitarianism as a normative view.18 This is reflected in the most import recent secondary literature.19

---

17 The passages from the *Analogy of Religion* are not in the Selby-Bigge anthology since when the anthology was created the *Analogy* it was no longer obviously a work of philosophy or relevant to philosophy.
18 For example Broad (1930: 81).
19 For example Penelhum refers to ONV as “Butler’s anti-utilitarian arguments” (Penelhum 1986: 82-85) as does Irwin (2008: 525). Both of these authors are (obviously) immensely learned, deeply immersed in all of Butler’s works, and offer
Whewell and Sidgwick differed in their assessment of the respective merits of these normative theories and the harm of the fifth objection to utilitarianism. But for all, the fifth objection highlighted the conflict between two normative theories. I would like to suggest there is a different way of reading it. Butler is not focusing on intuitionism as superior to utilitarianism, but rather on the idea that nothing derived solely from theoretical reason is exclusively the whole of virtue.

Now let’s return to Rawls’ description. It is mainly accurate if – as might be expected from lecture notes – not particularly edifying. But there is a subtle point of interpretation nonetheless which falls in line with the others described above. Rawls’ phrase “Butler entertains the purely hypothetical possibility that God might follow the principle of greatest happiness” implies that God does not follow the principle of greatest happiness. This would suggest that the conflict Butler is interested in is between the greatest happiness principle and what our ordinary moral faculties tell us, i.e. a normative conflict.

Butler does not use the phrase “greatest happiness” which although to be found in Hutcheson, has from Bentham on has become associated with utilitarianism as a normative theory (and thus signals that the conflict is one between normative theories). The stress in Butler’s passage, though, is on “nothing to himself as an end other than happiness – i.e. the exclusive nature of the explanation. In other words it is not on greatest happiness – it is on solely happiness. Consequently what Butler seems to be suggesting is that were God only interested in happiness, this would make what we take commonly to be virtue – being specially benevolent to one’s children – to be vice. But this is of course reversible. If God were only interested in righteous punishment, then the desire to happiness might be vice. Butler is less interested in a conflict between normative theories than in the inevitability of self-refutation in exclusive moral justifications.

I would suggest that there is a further reason why Butler might not be focusing on the conflict between utilitarian and deontic theories. That there are a few exclusive normative theories, or that normative accounts were independent of meta-ethics or moral epistemology, was not yet in currency because exclusive normative views were the exception not the rule. Hutcheson’s view was mixed in ways pointed out by Gay. The default position was a pluralism that invoked virtue in many contexts (see Gill 2014). Consequently it makes more sense to view the passage as attacking exclusive theories, the main examples of which was John Gay. But before John Gay it sophisticated, sensitive readings that allow for the ambiguous context. They are not reading from Selby-Bigge! My point is just that the backstory makes it desirable to cast what Butler is doing as an “anti-utilitarian argument”.

20 Rawls goes on to make the very edifying (and I think correct point) that there is continuity between Butler and Hume here despite Hume’s hostility to philosophical theology.
is very hard to give an example of a consistent, exclusive normative theory other than divine command theories.\textsuperscript{21}

This gets to the aforementioned Clarke. There is a second point being made here, that one ought not imitate God in deciding what is one’s best conduct. Clarke allowed different guides to human moral conduct including equity and benevolence (Clarke 1706: 82 ff.). It was unproblematic to Clarke that there are non-exclusive guides to conduct for a similar reason that it probably was to Butler, assumptions that love of God trumped and providence worked out the details. For Clarke the existence of God was deducible and the providentialism that followed from this certain. So Butler shared Clarke’s belief in plural duties or values anchored by divine governance.

But, as noted, Butler and Clarke differed on the best method in ethics, and we can see in this passage the difference coming to a head in a way that it didn’t in the \textit{Sermons}.\textsuperscript{22} Clarke stresses that we ought to maximize the benevolence in our action because God is maximally benevolent and so maximizing benevolence provides a guide to our conduct (Clarke 1706: 92-97). But, as Butler is noting in the fifth objection, what we know about our frames conflicts with taking a general standard that applies to God as a guide to our conduct, even if in a finite measure. Clarke is alluded to\textsuperscript{23} in a passage in I.vi.12 very close (i.e. two paragraphs before) the paragraph that Butler references in the fifth objection.

In the main body of the text Butler notes:

\begin{quote}
For the Conclusion, that God will finally reward the righteous and punish the wicked, is not here drawn, from it appearing to us fit, that He should; but from it appearing, that He has told us, \textit{He will}. And this he hath certainly told us, in the Promise, and Threatning, which it hath observed the Notion of a Command implies, and the Sense of Good and ill Desert which he has given us, more distinctly expresses. And this Reasoning from Fact is confirmed, and in some degree verified, by other facts. (Butler 1736: 169-170)
\end{quote}

\begin{footnotesize}
\begin{footnote}
I have much more to say about this but cannot within the restricted context of this essay. In short I hold that Gay really initiated the distinction between normative explanation and moral epistemological (moral sense) and meta-ethical considerations, and before Gay these distinctions were unclear.
\end{footnote}
\begin{footnote}
Sidgwick noted a shift in Butler from the \textit{Sermons} to ONV (1906: 86n) to recognizing the conflict between utilitarianism and ordinary intuitionist notions of virtue. This is why ONV is so crucial for Sidgwick.
\end{footnote}
\begin{footnote}
That Butler has the \textit{Sermons} distinction between Clarkean reasoning about fitnesses and his own reasoning from fact in mind is underscored by a more technical footnote that makes stronger use of Clarke’s language (Ibid169n1). Butler does love technical footnotes!
\end{footnote}
\end{footnotesize}
When we ignore reasoning from matters of fact and rely entirely on reasoning from abstract relations, this allows self-deceit in for a very particular reason. In the “Introduction” to the Analogy, Butler remarked that

Forming our Notions upon Reasoning from Principles which are certain, but applied to Cases to which we have no ground to apply them, like those who explain the Structure of the human Body, and the Nature of all Diseases and Medicine from mere Mathematicks; is an Error much a-kin to the former; since what is assumed to maker the Reasoning applicable, is Hypothesis. But it is allowed just, to join abstract Reasonings with the observation of Facts, and to argue from such Facts as are known, to others that are like them. (Butler 1735: v).

When reasoning is secure but there are no precise rule by which the reasoning is to be applied then as in Cartesian physiology another hypothesis is often tendered. In moral cases such as Clarke’s benevolence rule or Gay’s happiness, when the agent has a great deal of leeway the tendering of further hypotheses allows self-deception in in the choice of how to make the rule fit the indeterminate circumstances on the ground.\(^{24}\) I suggest that he is making a similar point in the fifth observation, a point that holds independent of particular normative accounts. It seems likely the purpose of the passage is to convince readers to avoid philosophies that promote “benevolent” from abstract and exclusive principles (exclusive in this case of what might being told us by conscience) insofar as – “such supposed endeavors to promote happiness for many others proceed, almost always, from ambition, the spirit of party, or some indirect principle, concealed perhaps in great measure from persons themselves” (see Appendix). In other words the fifth objection condemns philosophies that promote self-deceit by arguing that we ought to pay attention to general metaphysical reasons (or fitnesses) for conduct over and above the reasoning from fact.

That this is what Butler has in mind is reinforced by the passages from the Analogy to which he refers in the fifth objection. The main focus of Analogy I.vi – “Of the Opinion of Necessity considered influencing Practice” concerns how general necessitarian metaphysical doctrines offer poor and self-refuting guides to conduct; “with regard to Practice, it is as if it were false, so far as Experience reaches” (Butler 162). I conclude that Butler is making a general methodological point concerning theories that draw exclusively on abstract relations, are self-refuting in practice, and as a consequence are particularly prone to self-deception, not a point about combatant normative theories.\(^{25}\) Although Butler held that reasoning from abstract

---

\(^{24}\) Butler discusses this sort of self-deceit extensively in Sermon X where he discusses the example of Balaam. See Garrett 2014.

\(^{25}\) Butler adds “such is in Fact our Condition and the natural Course of things, that whatever we apply it to Life and Practice, this Application of it, always, misleads us, and cannot but mislead us, in a most dreadful Manner” (Ibid.). Self-deception is stressed by Penelhum in his interpretation (1986: 82-5).
relations and from matters of fact were compatible, I suggest that after Gay\textsuperscript{26} he became more worried about Clarke’s influence on abstract exclusivist theories.\textsuperscript{27}

So to conclude this discussion it seems from the evidence of who and what Butler is likely criticizing and the context of the passage in Butler’s longer philosophical works (i.e. the \textit{Analogy}) that the purpose of the discussion is not to highlight a conflict between normative theories but instead to highlight problems of methodological exclusivity – excluding matters of fact due to method and relying only on abstract reasoning and as a consequence excluding other sorts of moral principles. This is the methodological analogue of Butler’s stress on balance, moderation, and the importance of a multiplicity of principles (Schneewind 1998: 342-5) and his commitment to the plurality of values throughout (Gill 2014). Exclusive reliance on one moral justification results in ignoring or conflicting with another. Exclusive reliance on speculative reason results in allowing self-deceit in via the ignoring of matters of fact through our self-deceit in deciding how the reasons fit the facts.

This is not to deny that the standard themes might be there, but rather to argue that the themes I have just discussed are much harder to the emblematic themes of the standard interpretation. Consequently, it’s very difficult to look at the passage and not to see it as about utilitarianism. Emblematic passages are often like this, they can be taken as insightful for a cluster of different views, and what they are emblematic of is a function of how the insights are prioritized. In this case the passage (and Butler) has come to stand for something quite particular and important, and that it has become emblematic retrospectively reinforces this reading via anthologies, its place in the curriculum, and the persistence of the conflict in which it is emblematic. As suggested in the first part of the essay, I wish to underscore that due to the background history of the text (the anthologies, editing, curricular place, etc.), the special character of ethics, the emblematic place of the

\textsuperscript{26} One worry is that if Gay is the target, and Gay is the first consistent utilitarian, that Butler is attacking utilitarianism. In response I would underscore that since Gay is the sole example of a cogent exclusive normative philosophical theory (Calvinists were cogent but not philosophical in this sense) it makes it unlikely that Butler is attacking conflicts between normative theories.

\textsuperscript{27} In further support of this reading of the fifth objection, Butler remarks “For to pretend to act upon Reason, in Opposition to practical Principles, which the Author of our nature gave us to act upon; and to pretend to apply our Reason to Subjects, with regard to which, our own short Views, and even our Experience, will shew us, it cannot be depended upon” (Ibid. 164). The paragraph that Butler explicitly references stresses that there is nothing wrong with speculative reason in and of itself but rather our capacity for self-deceit ought to make us continually wary of how its application allows in “Prejudice and Perversion”. Butler is furthermore worrying, like Berkeley and many others did, about how speculation of this sort can provide a warrant to immoral actions and self-deception by reformers and others who are not philosophers.
passage in arguments by influential philosophers, and present concerns, there will much less motivation to read the text (and texts in general) in ways that are less “philosophically” pressing however plausible the interpretation. Indeed it seems as if there are further aspects of the passage, the focus on self-deceit and the practices of reformers, that are to us barely recognizably philosophical and we push past them to get to the core of the passage. Why and how this is the “core” is of course the issue. The purpose of my discussion though is not to say one ought only to read it this way, if I am correct then I am guided by present concerns as well,\(^{28}\) but rather to exhibit an emblematic passage and show how it is read, and – too sketchily -- why.

REFERENCES


---

\(^{28}\) Maybe they are now easier to see because of the rise of the post-Wolf and post-Williams interest in non-exclusive normative theories.


APPENDIX -- The fifth objection

"Fifthly, without inquiring how far, and in what sense, virtue is resolvable into benevolence, and vice into want of it; it may be proper to observe, that benevolence, and the want of it, singly considered, are in no sort the whole of virtue and vice. For if this were the case, in the review of one's character, or that of others, our moral understanding and moral sense would be indifferent to every thing, but the degrees in which benevolence prevailed, and the degrees in which it was wanting. That is, we should never approve of benevolence to some persons rather than others, nor disapprove injustice and falsehood upon any other account, than merely as an overbalance of happiness was foreseen to be likely produced by the first, and of misery by the second. But now, on the contrary, suppose two men competitors for any thing whatever, which would be of equal advantage to each of them; though nothing indeed would be more impediment, than for a stranger to busy himself to get one of them preferred to the other; yet such endeavor would be virtue, in behalf of a friend or a benefactor, abstracted from all consideration of distant consequences: as that example of gratitude, and cultivation of friendship, would be of general good to the world. Again, suppose one man should, by fraud or violence take from another the fruit of his labor with intent to give it to a third, who, he thought, would have such pleasure from it as would balance the pleasure which the first possessor would have had in the enjoyment, and his vexation in the loss of it: suppose also, that no bad consequences would follow; yet such an action would surely be vicious. Nay, farther, were treachery, violence, injustice, no otherwise vicious, than as foreseen likely to produce an overbalance of misery to society; then,

29 The first paragraph is my transcription from Butler 1736. The remaining paragraphs are thanks to David McNaughton and from his transcription.
if in any case a man would procure to himself as great advantage by any act of injustice, as the whole foreseen inconvenience, likely to be brought upon others by it, would amount to, such a piece of injustice would not be faulty or vicious at all; because it would be no more than, in any other case, for a man to prefer his own satisfaction to another’s in equal degrees. The fact then appears to be, that we are constituted so as to condemn falsehood, unprovoked violence, injustice, and to approve of benevolence to some preferably to others, abstracted from all consideration which conduct is likeliest to produce an overbalance of happiness or misery. And therefore, were the author of nature to propose nothing to himself as an end but the production of happiness, were his moral character merely that of benevolence; yet ours is not so. Upon that supposition, indeed, the only reason of his giving us the above-mentioned approbation of benevolence to some persons rather than others, and disapprobation of falsehood, unprovoked violence, injustice, must be vice in us, and benevolence to some preferably to others, virtue, from all consideration of the overbalance of evil or good which they may appear likely to produce. ONV §8

[The passage is followed by Butler’s reflection on the 5th observation, which includes the passage from Sidwick mentioned above. -- AG]

Now if human creatures are endued with such a moral nature as we have been explaining, or with a moral faculty, the natural object of which is actions: moral government must consist, in rendering them happy and unhappy, in rewarding and punishing them, as they follow, neglect, or depart from, the moral rule of action interwoven in their nature, or suggested and enforced by this moral faculty; in rewarding and punishing them upon account of their so doing.

I am not sensible, that I have, in this fifth observation, contradicted what any author designed to assert. But some of great and distinguished merit,* have, I think, expressed themselves in a manner, which may occasion some danger, to careless readers, of imagining the whole of virtue to consist in singly aiming, according to the best of their judgment, at promoting the happiness of mankind in the present state; and the whole of vice, in doing what they foresee, or might foresee, is likely to produce an overbalance of unhappiness in it: than which mistakes, none can be conceived more terrible. For it is certain, that some of the most shocking instances of injustice, adultery, murder, perjury, and even of persecution, may, in many supposable cases, not have the appearance of being likely to produce an overbalance of misery in the present state; perhaps sometimes may have the contrary appearance. For this reflection might easily be carried on, but I forbear — the happiness of the world is the concern of him, who is the Lord and the Proprietor of it: nor do we know what we are about, when we endeavour to promote the good of mankind in any ways, but those which he has directed; that is indeed in all ways, not contrary to veracity and justice. I speak thus upon supposition of persons really endeavou ring, in some sort, to do good without regard to these. But the truth seems to be, that such supposed endeavours proceed, almost always, from ambition, the

---

30 A l. vi. 14.
spirit of party, or some indirect principle, concealed perhaps in great measure from persons themselves. And though it is our business and our duty to endeavour, within the bounds of veracity and justice, to contribute to the ease, convenience, and even cheerfulness and diversion of our fellow-creatures: yet from our short views, it is greatly uncertain, whether this endeavour will in particular instances, produce an overbalance of happiness upon the whole; since so many and distant things must come into the account. And that which makes it our duty, is, that there is some appearance that it will, and no positive appearance sufficient to balance this, on the contrary side; and also, that such benevolent endeavour is a cultivation of that most excellent of all virtuous principles, the active principle of benevolence.

However, though veracity, as well as justice, is to be our rule of life; it must be added, otherwise a snare will be laid in the way of some plain men, that the use of common forms of speech generally understood, cannot be falsehood; and, in general, that there can be no designed falsehood without designing to deceive. It must likewise be observed, that in numberless cases, a man may be under the strictest obligations to what he foresees will deceive, without his intending it. For it is impossible not to foresee, that the words and actions of men in different ranks and employments, and of different educations, will perpetually be mistaken by each other: and it cannot but be so, whilst they will judge with the utmost carelessness, as they daily do, of what they are not, perhaps, enough informed to be competent judges of, even though they considered it with great attention. (ONV §§9-11)