

**BELGRADE PHILOSOPHICAL ANNUAL/  
FILOZOFSKI GODIŠNJAK 34/2021**  
Institute of Philosophy, Faculty of Philosophy,  
University of Belgrade  
Belgrade, Čika Ljubina 18–20

**Belgrade**

**Year XXXIII**

**YU ISSN 0353-3891**

**UDK-1**

### **Editor**

Slobodan Perović (University of Belgrade)

### **Associate Editors**

Miloš Arsenijević (University of Belgrade)

Jovan Babić (University of Belgrade)

Leon Kojen (University of Belgrade)

Živan Lazović (University of Belgrade)

Timothy Williamson (University of Oxford)

### **Editorial Board**

Berit Brogaard (University of Missouri, St. Louis)

Paul Boghossian (NY University)

Aleksandar Jokić (Portland State University)

Jan Narveson (University of Waterloo)

Georg Meggle (University of Leipzig)

J. Angelo Corlett (San Diego State University)

Howard Robinson (Central European University)

### **Managing Editor**

Petar Nurkić (University of Belgrade)

petar.nurkic@f.bg.ac.rs

Belgrade Philosophical Annual online at <http://www.f.bg.ac.rs/bpa/index.html>

**Printed by**  
Službeni glasnik, Belgrade

**This issue is financially supported by**  
Ministry of Education, Science and Technological Development of the  
Republic of Serbia

The statement on publication ethics can be found at the journal website  
<http://www.f.bg.ac.rs/bpa>

BELGRADE PHILOSOPHICAL ANNUAL 34/2021

HUME'S LEGACY

*Guest editors:* Peter Kail, University of Oxford

Angela Coventry, Portland State University

Dejan Šimković, The University of Notre Dame, Australia

Tamás Demeter	Towards a Humean Epistemic Ideal: Contested Alternatives and the Ideology of Modern Science.....	7
Hsueh Qu	Hume and Reliabilism .....	27
Sofía Calvente	The B Side of Imagination. Hume on Imperfect Ideas .....	53
Gabriel Watts	Reading Hume on the Passions.....	73
Jonas Olson	Hume's Sentimentalism: Not Non-Cognitivism ....	95
Aleksandra Davidović	Who Speaks for Hume: Hume's Presence in the <i>Dialogues Concerning Natural Religion</i> .....	113
Liz Goodnick	The Role of the <i>Natural History of Religion</i> in Hume's Critique of Religious Belief .....	139
Emilio Mazza	As "men of sense". Godwin, Baroja, Bateson and Hume's "Of National Characters" .....	159



# HUME'S LEGACY



## TOWARDS A HUMEAN EPISTEMIC IDEAL: CONTESTED ALTERNATIVES AND THE IDEOLOGY OF MODERN SCIENCE<sup>1</sup>

**Abstract:** *I suggest that it is fruitful to read Hume's Enquiry concerning Human Understanding as a concise exposition of an epistemic ideal whose complex philosophical background is laid down in A Treatise of Human Nature. Accordingly, the Treatise offers a theory of cognitive and affective capacities, which serves in the Enquiry as the foundation for a critique of chimerical epistemic ideals, and the development of an alternative ideal. Taking the "mental geography" of the Treatise as his starting point, this is the project Hume pursues in the Enquiry. The epistemic ideal Hume spells out in the Enquiry is an alternative to competing ideals: the Aristotelian, the Cartesian, and the Newtonian, and can be read as an exposition of the epistemic ideal of modern science. Although the spell of the Aristotelian and the Cartesian ideals had been in decline for several decades by the 1740s, they had not fully lost their grip on the philosophical imagination. Yet, it was the Newtonian epistemic ideal that became dominant in Scotland and Britain by then, guiding inquiry in moral and natural philosophy, as well as in medical theory. Hume offers a critique of these ideals. He shows that Aristotelian and Cartesian epistemic aspirations rest on mistaken views on human cognitive capacities. And albeit the Newtonian ideal is not prone to this mistake by Hume's standards, its epistemic expectations extend far beyond the limits of those capacities. Hume's epistemic ideal can be read as a correction, limitation and refinement of the Newtonian ideal: it sets epistemic aims and propagates methods for the production of fallible, limited and potentially useful knowledge that falls short of the great epistemic expectations of Newton and many Newtonians – but it conforms to what we expect from modern science.*

**Keywords:** *Hume, epistemic ideals, scientia, demonstrativity, Aristotle, Descartes, Newton*

### 1. Introduction

Robert Pasnau (2013, 2017) has argued inspiringly that historically speaking, epistemology had been centred on the exposition of *epistemic ideals* – and the focus on the conceptual analysis of 'knowledge' would gain

---

1 This paper contributes to the research programme of MTA Morals and Science Lendület Research Group. I am grateful for comments and suggestions to the members of the group, particularly Gábor Biró, Péter Hartl, Gergely Kertész, Gábor Zemplén and Deodáth Zuh.

prominence only in the twentieth century, so the history of philosophy teaches that “a theory of knowledge ought not to be a theory of ‘knowledge’” (Pasnau 2017, 11). Epistemologies aiming at the development or refinement of epistemic ideals have not been preoccupied with the necessary and sufficient conditions to meet when we apply ‘knowledge’ to some epistemic state of persons: if ‘knowledge’ is taken to express an epistemic ideal that inquiry should approach, then the ideal itself may remain unattainable and yet orient successfully towards desirable aims and methods. Aristotle’s own writings, for example, “contain no examples that satisfy all the necessary requirements” of *episteme* (*ibid.* 5) – an ideal whose central tenets had been dominating epistemological thinking up until about the eighteenth century. So, epistemic ideals are proposed with an eye to what our inquiries should aspire to – even if it is impossible to achieve. Just like conceptual analysis, constructing epistemic ideals is also a normative enterprise, but in a different way: they do not inspire to demarcate cases of knowledge from non-knowledge, but to provide motivation and guidelines for inquiry. An epistemic ideal shows what inquiry should achieve and how it should achieve it – never mind that “we commonly fall off from that ideal” (*ibid.* 3).

Epistemic ideals can be developed only as part of complex philosophical systems, because they presuppose a metaphysical apparatus and a theory of cognitive capacities. In order to formulate an epistemic ideal, one needs a theory of what can be known at all – and one has to answer questions about how, and to what extent, our cognitive capacities are suited to provide access to those possible objects of knowledge. Plato’s epistemology would be incomprehensible without his theory of Forms, Aristotle’s *episteme* presupposes his essentialism, and Cartesian science rests on the foundations of a metaphysics of modes, etc. So, an “epistemic ideal depends critically on how one conceives of the world” (*ibid.* 58). This account of what the world consists of has to be accompanied by a theory of perception and reasoning – the two ways human beings can gain epistemic access to the world. How brave our epistemic expectations can be, how their prospects and limits could be articulated depend on these cognitive and metaphysical considerations.

Here I am going to suggest that it is fruitful to read Hume’s *Enquiry concerning Human Understanding* (1748, hereafter EHU) as a concise exposition of an epistemic ideal whose complex philosophical background is laid down in *A Treatise of Human Nature* (1739/40, hereafter T). As I do not have the space here to dwell on the background the *Treatise* provides, I only indicate that I consider this suggestion consonant with Hsueh Qu’s recent work. Qu (2020), rightly I think, points out a crucial difference between the projects of the *Treatise* and the *Enquiry*: epistemological issues in the *Treatise* tend to arise “as offshoots of psychological discussions” (*ibid.* 5), while the *Enquiry* is “much more epistemological in nature” (*ibid.* 6). The fundamental ambition of the *Treatise* is “to conduct the study of human nature as a science”



(*ibid.* 41), while in the *Enquiry* “this philosophical undertaking is not taken as an end in itself, but it is rather instrumental towards the epistemic goal of destroying false philosophy” (*ibid.* 43). This latter project can be fruitfully explained in terms of epistemic ideals. The *Treatise* offers a theory of our cognitive and affective capacities, which serves in the *Enquiry* as a foundation for a critique of chimerical epistemic ideals, and the development of an alternative ideal. Taking a “mental geography” as his starting point (EHU 1.13), this is the project Hume pursues in the *Enquiry*.

The epistemic ideal Hume spells out in the *Enquiry* is an alternative to *competing ideals*: the Aristotelian, the Cartesian, and the Newtonian, and can be read as an exposition of an epistemic ideal of modern science. The spell of the Aristotelian and the Cartesian ideals had been in waning for several decades by the 1740s, they had not fully lost their grip on philosophical imagination. Yet, it was the Newtonian epistemic ideal that became the dominant force by then, guiding inquiry in moral and natural philosophy, as well as in medical theory.<sup>2</sup> Hume offers a critique of these ideals. He shows that Aristotelian and Cartesian epistemic aspirations rest on mistaken views on human cognitive capacities. And although the Newtonian ideal comes much closer to Hume’s standards, it also tends to extend its epistemic expectations far beyond the limits of those capacities. Hume’s epistemic ideal is thus a correction, limitation and refinement of the Newtonian ideal: it sets epistemic aims and propagates methods for the production of fallible, limited and potentially useful knowledge that falls short of the great epistemic expectations of Newton and many Newtonians – but it conforms to what we expect from modern science.

In the context of competing epistemic ideals, the project of the *Enquiry* can be framed with an eye on contemporaneous *genres of knowledge production* and the *epistemic concepts* integrating them. This philosophical project begins with clarifying crucial epistemic concepts such as experience, reasoning, inference, induction, probability, causation, explanation and testimony. In this respect, the aim of the *Enquiry* is to chase out “obscurity in the profound and abstract philosophy” (EHU 1.11) by the “spirit of accuracy” (EHU 1.9). The nature of “accurate and just reasoning” to be deployed in these matters can be revealed only “from an exact analysis of [our cognitive] powers and capacity” (EHU 1.12). The *Enquiry*, however, does not provide such an analysis but rather invokes its results: the task of analysis had already been taken up in the *Treatise* and produced an “anatomy” of the mind (T T 1.4.6.23, 2.1.5.6 and 8, 2.2.11.6, 2.1.12.2, 3.3.6.6, A.2, see also EHU 1.8, 7.9).<sup>3</sup> Against this background the *Enquiry* is left with the task of subverting “abstruse philosophy and metaphysical jargon” and “cultivat[ing] true

2 I discuss several aspects of this situation in Demeter 2016.

3 I sketch the character of this anatomy in Demeter 2016 ch.8 and more recently in Demeter 2021.

metaphysics” (EHU 1.12) by revealing the content that epistemic concepts can legitimately have, given their genealogy in human nature.

These epistemic concepts play a foundational and integrative role: epistemic practices, from observation to theory construction, are organized around them. Exposing their “distinct meaning”, “nature” and “reality” (as suggested in EHU 2.9) reveals the kinds and the strength of knowledge claims that they can support in various genres of knowledge production, and thereby it reveals what kind of epistemic benefit we can hope from the practices organized around them. Hume’s analyses define an epistemic ideal: they show what our expectations should be concerning our evidential, inferential, explanatory and predictive practices as integrated by these concepts. Relying on an empiricist genealogy of these concepts, Hume outlines what we can expect from institutionalized genres of knowledge production: natural and moral philosophy, mathematics, and theology. Following a critique of alternative epistemic ideals and a critique of epistemic concepts, a critique of the genres of knowledge production constitutes the third angle from which Hume’s epistemic ideal can be characterized.

This paper focuses only on the first step of this programme for reconstructing Hume’s epistemic ideal, namely the alternative stances against which Hume’s ideal is unfolding. Locating Hume’s epistemic ideal in the proper historical context in this way is important, because the elaboration of epistemic ideals is a cooperative process that brings about advances in metaphysics and epistemology and results in a refined and complex epistemic ideal. But Hume’s *Enquiry* is rather concise, and we should not expect detailed critiques of competing ideals in it. Hume’s challenge is advanced only against some of their core elements.

## 2. The Aristotelian ideal

Historically, Aristotle’s *episteme*, subsequently known as *scientia*, was unquestionably the most influential and persistent epistemic ideal, several tenets of which survived even in the works of those criticizing it.<sup>4</sup> This ideal has its roots in Aristotelian logic: according to Aristotle’s *Posterior Analytics*, knowledge worthy of its name must consist of propositions that are *universally and necessarily true*. This necessity must be demonstrated through *sylogistic inferences* that proceed from premises containing the *cause* of the conclusion. This procedure leads to *certain knowledge* and *understanding* of the *necessary cause* of the phenomena and an *explanation* of why it is the way it is and cannot be otherwise. Being necessary and universal, this knowledge is focused on the *essential structure* of things, and through their essences it accounts not only for how things actually behave, but also for how they would behave in a variety of circumstances. Accordingly, knowledge or *scientia* consists in a

4 See Pasnau (2017), and for a more concise overview see Demeter, Láng, Schmal (2016).

*systematic, demonstrative* presentation of why things behave the way they do – and not in the discovery of the causes from which demonstrations followed.

Even if this ideal had never been fully implemented in epistemic practice (not even by Aristotle himself, see Pasnau 2013, 993), it served as the guide of systematic inquiry for several centuries. Particularly, Aristotelianism had been a definitive force in Scottish philosophy well into the eighteenth century (Sgarbi 2013, 44–51), and the Aristotelian ideal was not yet a thing of the past in Hume’s time. While studying in Edinburgh, Hume might have been exposed to the Aristotelian ideal in Colin Drummond’s class centred on syllogism and Aristotelian metaphysics, which “gave Hume an edifice to destroy, not a path to follow” (Harris 2015, 39). And the influence of Aristotelian logic of inquiry extended beyond the time of Hume’s education in the 1720s: Thomas Reid in 1774 still considered it a worthwhile undertaking to point out the many features of Aristotle’s logic that prevent it from being a useful “engine of science” (Reid 1872, 1:701–702).

*Repeated experience* has a crucial role to play in the Aristotelian ideal: it is supposed to provide access to the essence of phenomena from which demonstrations can follow. At this point Aristotelian method and metaphysics are combined, so the ideal of *episteme* entails essentialism. Although Hume is not an enemy of experimental inquiry into the hidden compositions and structures,<sup>5</sup> he also holds that the “ultimate springs and principles are totally shut up from human curiosity and enquiry” (EHU 4.12), and that no amount of experience can help us in this regard. With the help of experience we can “resolve the many particular effects into a few general causes” (EHU 4.12), and that by such causal reasoning we can attain some “assurance concerning objects, which are removed from the present testimony of our memory and senses” (EHU 7.29). Consequently, our inductive practices lend some credibility to our inferences concerning “secret powers” (EHU 4.21) that might be resolved further by the help of experience, but even the “most perfect philosophy of the natural kind only staves off our ignorance a little longer” (EHU 4.12).<sup>6</sup> So, no knowledge of Aristotelian essences is to be hoped for.

As for Aristotelian methodology, it is less than ideally clear how the step from particular experiences to universal essences is to be taken, but the *Organon* provides at least two possibilities for understanding Aristotelian induction. One way is to claim that the inductive step is *intuitive*, so that “from several particulars the universal is clear” (88a4 see also 100b4–5). Induction so understood is “a natural psychological process by which intuition of self-evident truths develops in us over time” (Salmieri 2014, 3). The other possibility (suggested e.g. in 156b10–16) is to take induction

---

5 Different versions of what Humean inquiry into hidden structures means are presented by Landy (2018) and Demeter (2020).

6 Elsewhere I argued that these passages in Hume can be fruitful to read as expressing commitment to a version of constructive empiricism (see Demeter 2020).

as an argument from the particulars to the universal (see also Smith 2017, 3.1). Either way, as a result of induction we get a view of the intelligible form beyond the perceptual form of phenomena – i.e. a view beyond what is given in experience, and it reveals to us *what it is* that is given in experience (see also Buckle 2001, 91). This view is an epistemic state that Aristotle calls *nous*, the apprehension of essences as the causes from which phenomena can be demonstrated as necessary effects arising from necessary causes. This demonstrative knowledge constitutes *episteme* or *scientia*.

Hume's *Enquiry* can be read as arguing that the epistemic burden repeated experience is supposed to carry in the Aristotelian ideal is much bigger than on philosophical scrutiny it can bear. In Section 4 Hume undermines the suggestion that “uniform experience” can lead to knowledge through reasoning and reflective inference. In the form of a poetic question whose answer he cannot even imagine, he points out that there is no “process of reasoning, which, from one instance, draws a conclusion, so different from that which it infers from a hundred instances, that are nowise different from that single one” (EHU 4.20, see also 4.21). Repetition and uniformity of experience can only lead to the conclusion that similar causes have “always been attended with” similar effects, but they cannot establish universal conclusions that the same causes will, let alone must, be attended with the same effects – as Aristotelian *scientia* would require. There is no intuitive step leading from “always” to “will” or “must”, nor to the missing intermediate steps between them for drawing the inference needed (EHU 4.16).

While in the Aristotelian ideal repeated experience is supposed to lead to an epistemic state (*nous*) by intuition or reasoning, in Hume's alternative picture the immediate yield of repeated experience is only an “instinct or mechanical tendency” distinct from our reflective inferential capacities (EHU 5.22). This tendency, namely custom, brings along a doxastic state that is “felt by the mind” (EHU 5.12) – i.e. an affective rather than an intellectual state of mind. The mechanical tendency inculcated by repeated experience can only give rise to an impression of reflection, but no intuition: it can only yield an expectation that directs the course of our ideas in our inferences concerning factual matters. Consequently, the idea of cause turns out to be a sense-based concept in Hume's hands (Garrett 2015, 129–136).

The doxastic state arising due to repeated experience is much more malleable than *scientia*. It is not related to essences, cannot aspire to universality and certainty, but it is open to revision: the course of our experience can change, and even more likely, new experience can change the doxastic state itself. Hume's account of what repeated experience can deliver is thus more congruent with the Newtonian ideal of open-ended inquiry where experimental knowledge is the basis of further inquiry – and not the foundation of ultimate knowledge (see e.g. Smith 2002, 160–161).

### 3. The Cartesian ideal

The Cartesian epistemic ideal emerged as an influential and effective alternative to the Aristotelian. The Cartesian ideal diverged in several respects from the Aristotelian while preserving some of its characteristics, most notably its aspiration for *demonstrativity* and *certainty*. Descartes was still devoted to the ideal of *scientia* insofar as the aspiration to certainty is concerned. This certainty, however, was not supposed to be reached through demonstration by syllogistic inference that Descartes severely criticized as a useless logical instrument for delivering knowledge. Instead, he advocated a method of *analysis* and *synthesis*: breaking down complex questions into simple ones to which intuitively clear and distinct answers can be given and then to proceed from these intuitive answers to answering complex questions. The demonstrativity of this process is preserved by grasp of *intuitive* truths: each step in the process follows intuitively from the previous one. The reliability of this process is not due to the rules of syllogistic logic, but the benevolence of God, because it entails that God does not intend to deceive, and so it guarantees the reliable functioning of our faculties in the process of *cogitatio* (CSM 2:109–111).

The core of the Cartesian ideal is thus, similarly to the Aristotelian, a combination of metaphysics and method. Its metaphysics presented the world as consisting of two kinds of substance: matter whose essential attribute is extension, and whose properties are all modes of extension – and mind whose essential attribute is thinking, and whose properties are all modes of thinking. Its method is focused on the analysis of complex problems into simpler ones in a deductive manner until reaching those problems that are prone to “clear and distinct” answers by means of intuition – just like the most basic questions in mathematics. This method is supposed to guarantee the demonstrativity and certainty of the insights our inquiries can deliver.

The Cartesian method was focused on clear and distinct ideas, to be reached either by intuition or by the method of analysis, from which explanations could be constructed by synthesis. This method is formulated as Rule 5 for the direction of the mind:<sup>7</sup>

The whole method consists entirely in the ordering and arranging of the objects on which we must concentrate our mind’s eye if we are to discover some truth. We shall be following this method exactly if we first reduce complicated and obscure propositions step by step to

---

<sup>7</sup> Garber (2001, 35) denies this is a reference to analysis/synthesis, and he breaks with the tradition that treats these passages so. He would rather reserve analysis/synthesis as modes of theory construction. Smith (2010, 67, 84–86) is convenient to talk about the method advocated in the *Rules* in terms of analysis/synthesis. Jardine (1974, 249) gives a list of several terms that were used as synonymous in the period, and the wording of Rule 5 conforms to it.

simpler ones, and then, starting with the intuition of the simplest ones of all, try to ascend through the same steps to a knowledge of all the rest. (CSM 1: 20)

These two phases of method can come apart depending on the ideas involved in our inquiries. In geometry, for example, we can begin demonstrations in the phase of synthesis because they “deal only with the simplest and most general things” (CSM 2:14), so the “primary notions” here are “readily accepted by anyone” (CSM 2:111). Therefore, we can begin instantly with clear and distinct postulates, axioms, and definitions. In metaphysics, however, the “primary notions” are not at all clear in the first instance, so they must be excavated: they must be made clear and distinct by the means of analysis. And there are, of course, complex geometrical problems where synthesis can only be “a follow-up to analysis” (CSM 2:111; see e.g. the Pappus problem, Bos 2001, ch.23), and Descartes seems to suspect that even the clearest and most ancient examples of geometrical syntheses rested on analysis (CSM 2:111). So, both analysis and synthesis are geometrical methods of demonstration (CSM 2:110).

The Cartesian understanding of matter in terms of extension invites inquiry into nature to be conducted with geometrical methods. Consequently, it is natural to expect the same kind of knowledge from natural philosophy as from geometry. As Rule 5 above prescribes, inquiry should consist in the analysis of complex *propositions* into simple ones until reaching those that are susceptible of intuitive answers in terms of the primary objects of intuition: “simple natures”. Their class includes e.g. ‘cause’, ‘the one’, ‘the equal’, ‘thought’, ‘will’, ‘shape’, ‘extension’, ‘motion’ etc. and our acquaintance with them is “derived from the light of reason” (Garber 1992, 33). Garber (2001, 37) reconstructs a series of such reduction from the complex question of “What is the shape of a line (lens) that focuses parallel rays of light to the same point?” to “What is a natural power?” Once this question is answered by means of intuition we can “ascend through the same steps” answering eventually the initial question with as much clarity and certainty as the bottom question has been answered.<sup>8</sup>

This process of analysis results in a “hierarchical conceptual structure” (Smith 2010, 89) in whose discovery (in the phase of analysis) and deployment (in the phase of synthesis) experience has only an auxiliary role to play. While “we can use the connections we find in nature as a guide to the connections we seek in reason” (Garber 102), the epistemic burden is carried by deductions and clear and distinct intuition. So, experience can help structuring the process of reduction until we reach the unanalyzable question at the bottom that intuitively clear answer is to be given. Then in the phase of synthesis we can provide causal explanations by deriving observed phenomena from

<sup>8</sup> A more complex example is shown in the deduction of the cause of the rainbow, see Garber 2001, 100.

intuitively clear principles. As Garber (2001, 119) points out, experience here plays a role similar to that of diagrams in geometry: whatever actual visual experience, measurement and construction tell us about objects of geometry, geometrical knowledge is attainable only through the proof of a theorem. “Observation and experiment may play an important role in establishing an experimental fact, but it is reason that must confer the ultimate status of facthood on an observation” (Garber 2001, 313–314).

The Cartesian ideal did not enjoy a prolonged popularity in Scotland (see Shepherd 1982, Barfoot, 1990, Wilson 2009, 1–32). From the 1680s onwards, Gilbert McMurdo and Alexander Cockburn adopted Cartesian ideas, and they spread the mechanical worldview among their students, as was the case with most of their fellow regents in Edinburgh at that time. Until about 1690 Cartesianism prevailed in natural philosophy with a focus on the idea of rotating transparent matter causing the planets to orbit in the same direction and explaining why objects are falling toward the earth’s surface. After the publication of Newton’s *Principia*, forces quickly populated the mechanical universe and gravity replaced vortices in the explanation of planetary motions. The transformation of Cartesianism into Newtonianism took place in Edinburgh fairly rapidly, and by about 1710 the triumph of Newtonianism was eminent at the other Scottish universities as well.

Descartes was an important inspiration behind the construction of Hume’s *Treatise*. This should not come as a surprise as its composition took place mostly at the Royal College at La Flèche which “was a good place to read Descartes and to learn about Cartesianism” (Perinetti 2018, 65). In a letter dated 26/8/1737, Hume suggested to Michel Ramsay that reading Descartes’s *Meditations* might prove to be useful for understanding the metaphysical disquisitions in his then only forthcoming work (Norton and Norton 2007, 443). Descartes’s *Second Replies* (esp. CSM 2:110–111 discussed above) reveal how important the methodological lessons of the *Meditations* are. And arguably, the sharpest contrast between Descartes’s and Hume’s projects can be drawn in terms of method (see e.g. Wright 2009, 43), and this contrast is reflected in Hume’s normative epistemology.

At the beginning of Section IV Hume’s taking out the fork of the drawer sets the scene: by putting “matters of fact” and “relations of ideas” on two different epistemological plains (EHU 4.1), Hume exposes the gap on which the Cartesian methodology was built.<sup>9</sup> From Hume’s perspective Descartes’s deductive, and by Hume’s standards *a priori*,<sup>10</sup> method is untenable for inquiries into matters of fact, because those inquiries, as Hume keeps

9 Buckle (2001, 158–159), drawing on different pool of references disagrees with the claim that Descartes is the target here, but he also says: “Hume’s argument concerns a view that founds knowledge on veridical perceptions from which reason establishes a system of knowledge.” In the present reconstruction, this is the idea underlying Descartes’s method.

10 *A priori* in Descartes has different connotations from Hume’s and from the modern (post-Humean and post-Kantian) sense of the term, see Smith 2010, 67–69.

emphasizing, can only be directed by experience, observation, and *analogy* (EHU 4.12, 9.1, 5, 10.10, 11.27, 30,) – not by reason (EHU 4.7, 14). As all reasoning concerning matters of fact rely “on the relation of *Cause and Effect*” (EHU 4.4), and as this relation can be known exclusively from experience, it is hopeless to aspire to causal explanations that rely on deductions and intuitions. Cartesian “simple natures” cannot guide our inferences here, Hume because, according to Hume: “Solidity, extension, motion; these qualities are all compleat in themselves, and never point out any other event which may result from them” (EHU 7.8).

Consequently, the primary objects of inquiry must be those of experience and not of reason. Instead of resolving complex problems with methods akin to those of geometrical analysis, Hume proposes to “resolve the many particular effects into a few general causes, by means of reasonings from analogy, experience, and observation” (EHU 4.12). So, instead of *a priori*, geometrical approach to *propositions*, Hume advocates a comparative empirical approach to the *phenomena* themselves. These inquiries are, by the nature of experience, limited and fallible: they cannot produce ultimate and certain knowledge, but there is no other way at our disposal than the way of experience. And applying geometry in this field cannot be a remedy. For Descartes, geometry (analysis) plays the leading role in discovery, experience is only auxiliary; for Hume all the discovery of a matter of fact such as a law of nature “is owing merely to experience” and geometry can only “assist” us (EHU 4.13). *A priori* reasoning of any sort cannot help us in discovery, because then we could only consider an object “as it appears to the mind, independent of all observation” (EHU 4.13), and this consideration cannot lead us to anything distinct from the object – particularly it cannot suggest its effect or cause (EHU 4.11).

If despite all this, one still thinks it meaningful to undertake a Cartesian project of analysis and synthesis, then one fell prey to the appearance that custom induces: with respect to the familiar course of nature we imagine we can infer the effects “by the mere operation of our reason, without experience” (EHU 4.8). The Cartesian epistemic ideal is inspired by this deception rooted deeply in human nature – and given human nature, reason must give way to experience as the engine of inquiry.

#### 4. The Newtonian ideal

When Hume attended Edinburgh University in the early 1720s, the culture of inquiry was dominated by Newtonianism, and due to the work of David Gregory, John Keill and Colin Maclaurin, the influence of Scottish Newtonians extended well beyond the Scottish borders (see Wilson 2009, 34–59). Generally speaking, Hume was disappointed with the education he



received at the university and he had a very low opinion on the knowledge to be acquired there. One exception, at least to some extent, seems to be the natural philosophy class, which was taught to him by Robert Steuart (Stewart 2005, 25). In Steuart's class Hume presumably was required to study Keill's introductions to natural philosophy and astronomy, Gregory's introduction to optics and astronomy along with certain passages from Newton's *Opticks* and *Principia* (*ibid.* 21–22). Also, he probably made good use of the Physiological Library Steuart had established, which might have provided him with all the relevant literature he needed for an introduction to the problems of contemporary natural philosophy.

Newton's epistemic ideal went through significant changes, some of them due to the controversies arising in the aftermath of his new theory of light and colours, published in 1672 (see Zemplén and Demeter 2010). Initially, Newton was insisting on the idea that colour phenomena can be treated with mathematical certainty, and considered his famous *experimentum crucis* as outright demonstrative (Turnbull 1960, 80). On this basis he maintained that “*the Science of Colours will be granted Mathematicall & as certain as any part of Opticks*” (Turnbull 1959, 187). His confidence would be shaken subsequently. First, he retreated to the claim that optics operates with physical principles and mathematical demonstrations: physical principles constrain certainty, but if they are true, the conclusion of mathematical reasoning founded on them is demonstratively true (Turnbull 1959 p. 187–188). Then finally he proclaimed in Query 31 of the 1717 edition of the *Opticks* that “arguing from experiments and observations be no demonstration of general conclusions, yet it is the best way of arguing which the nature of things admits of” (Newton 2004, 138).

This can be read as an admission of failure to extend mathematical analysis at least to certain parts of nature – an admission that amounts to giving up the ambitions of the *Principia*. Newton's method in the *Principia* is focused on an analysis of motions and the forces producing them, and he makes it not less than “the basic problem of philosophy ... to discover the forces of nature from the phenomena of motions and then to demonstrate the other phenomena from these forces” (Newton 2004, 41). In the case of colours Newton had to give up his initial hopes for a demonstrative mathematical exposition that he had achieved for fits and refrangibility (Cohen 1980, 138–140). This might motivate Newton's permission for a potentially qualitative (as opposed to mathematical) analysis of “compounds to ingredients” as a route to explanatory principles in Query 31 of *Opticks*, and also his entertaining the possibility of accounting for optical phenomena as chemical phenomena (Shapiro 1993, 142n16). Eventually he had to allow for an experimental decomposition of white light into its component colours, but he had to stop short of giving it a full mathematical treatment in terms of motions and forces acting on light corpuscles (Guicciardini 2009, 316–317).

Hume's epistemic ideal is not consonant with the mathematically focused "strong program of mechanical natural philosophy" (Reill 2003, 27n14), a programme that, following Galileo's and Descartes's footsteps, was completed by Newton's *Principia* and had also been represented in his early optical inquiries. For Hume, mathematics is a useful tool in the practical applications of natural philosophy, but not in discovery: it is not possible to have a demonstrative, mathematical science about matters of fact, because the discovery of laws and our ability to draw causal inferences are owing exclusively to experience. So, "it is a law of motion, discovered by experience, that the moment or force of any body in motion is in the compound ratio or proportion of its solid contents and its velocity" and geometry can only assist us only when we are to *apply* this law e.g. to establish the proportions of mass and velocity when we are to move a great weight "by contrivance or machinery" (EHU 4.13, for a similar point see also T 2.3.3.2).

Given Hume's fork, this should not be surprising (see Slavov 2017). Natural philosophy, being concerned with matters of fact, cannot be based on mathematical *principles*, because they lack empirical content, so they cannot serve as the model of reality. Natural philosophy cannot apply mathematical *methods* of reasoning either, because "the only objects of the abstract science or of demonstration are quantity and number, and that all attempts to extend this more perfect species of knowledge beyond these bounds are mere sophistry and illusion" (EHU 12.27). So, mathematics cannot be a guide or model in our reasoning about matters of fact, and it cannot be seen as a guide for revealing the structure of the world. Mathematics just cannot be the meaningful language in which the book of nature is written – and this disqualifies the programme of the *Principia* as an epistemic ideal desirable by Hume's standards.

However, arguing from experiments and observations is, by Hume's standards, indeed the best (actually, the only) way if we aspire to epistemic benefits with respect to matters of fact – just as Newton proclaims in Query 31. Arguing this way, as is already clear (in EHU 4.12 discussed above), is arguing by analogies founded on similarities: "From causes, which appear *similar*, we expect similar effects. This is the sum of all our experimental conclusions" (EHU 4.20). Being founded on similarity, analogy also comes in degrees (see EHU 4.7–8, 4.20, 10.10 and esp. 9.1), and we can rely on it in our inferences only so far as similarity extends. Yet, we have a tendency to overstretch similarities and thus to violate the "rules of analogy" (EHU 11.27) by drawing inferences without being entitled to them.

This also means that the knowledge we can derive from such reasoning is fallible. There are uncertainties and irregularities in the course of events, so similarities can only lead us only to probable conclusions (EHU 6.4). Where no such irregularities occur, uniform experience can amount to "proof" (EHU 6.5) that leaves "no room for doubt or opposition" (EHU 6, n10). But even in these cases our certainty is not fallible, because it is not demonstratively

true that the course of nature will not change (EHU 4.18). Even experience combined with mathematics is still experience and arguments put forward this way (as in “mixed mathematics”) can only aspire to the status of “proof” and never to that of “demonstrations” (EHU 6, n10).<sup>11</sup> Striving for demonstrative knowledge of nature is thus doomed by the very nature of experience and experimental reasoning. This seems to be the conclusion Newton himself draw as a result of the controversy ensuing after his attempt at a demonstrative experimental science of colours (Zemplén and Demeter 2010, 651). This insight is also present in the way Hume draws and treats the distinction between proofs and demonstrations in the *Enquiry* (EHU 6, n10) – a distinction that is consonant with Newton’s typical usage.

Beyond the questions of demonstrativity and the applicability of mathematical methods to inquiry into matters of fact I take Hume’s epistemic ideal to be largely consonant with Newton’s in several respects (see Demeter 2019a).<sup>12</sup> The emphasis on what is actually observed is an expression of epistemic humility that was perceived as a Newtonian epistemic virtue and was expressed in various contexts, most famously perhaps with respect to the cause of gravity (Newton 2004, 100). Due to the limits of inductive methods the insights derived from the analysis of phenomena were taken to be subject to potentially open-ended further refinement, just as Newton envisaged (see Smith 2002, 160–161). So we can keep reducing our ignorance, but as Hume warns us, we should not hope to reveal ultimate knowledge of causes (EHU 4.12, see also e.g. T 1.2.5.25). They also agree that explanatory hypotheses unfounded on the analysis of experience are of no epistemic value. One consequence of the ban on explanatory hypotheses was the emphasis placed on explaining phenomena increasingly in their own terms as opposed to the speculative metaphysics of some underlying hypothetical reality (Gaukroger 2014, 20–28) – let it be the world of Aristotelian essences or that of allegedly clearly intuited Cartesian simple natures.

In Newton’s epistemic ideal speculative metaphysics is suppressed by what is actually observed and this tendency extends to metaphysical concepts: Newton transforms them, e.g. questions concerning space and God, into empirical questions of experimental natural philosophy (see Stein 2000, 261–262, 269–270, 277). This is what Schliesser (2011) calls “Newton’s challenge to philosophy”: a challenge that is posed to the concept of causation too: despite Newton’s famous denial in a letter of 1692 to Richard Bentley, he did believe in action at a distance (see Henry 2011), i.e. he relied on a concept of causation that did not require the spatial contiguity of cause and effect. This was a crucial metaphysical breach with the then dominant

11 Given the Hume’s account of the practices and metaphysics of mathematical knowledge, even the certainty and demonstrativity of mathematics can also be effectively questioned (see Demeter 2019b).

12 The literature on Hume’s relation to Newton is now voluminous. For the most recent overview see Schliesser and Demeter 2020.

concept of mechanical causation (see e.g. Ott, 2009) that allowed only for contact causation within the sphere of intelligible processes. The requirement of contiguity is still explicitly required for causation in the *Treatise* (see e.g. T 1.3.14.1, 1.3.15.3, Coventry 2006, 84–85), but it is not at all mentioned as a necessary condition of causal relation in the *Enquiry*.

Hume's analysis can be read as an "empirical deduction"<sup>13</sup> of a metaphysical concept in the spirit of Newton's challenge: a genealogy that reveals its traceable empirical content. Hume's billiard-ball example is mentioned only passingly in the *Treatise*, but it is thoroughly exploited and becomes crucial in the *Enquiry* (T 1.13.14.18, Abs.9, and EHU 4.8–10, 5.11, 7.6, 21, 28, 30). Its analysis, and the emphasis Hume puts on it, in the *Enquiry* supports the removal of the contiguity condition, because the crucial lesson that one can draw from the analysis of the experience of collision is that contact causation is not any more intelligible than action at a distance. In the *Enquiry*, constant conjunction and the ensuing idea of necessary connection are at the heart of Hume's analysis of 'cause' and it can accommodate Newtonian gravity as legitimate in causal explanations – even without specifying contact mechanisms by which it exerts its influence. This legitimacy cannot be provided by the *Treatise's* account.

Beyond these methodological and metaphysical convergences and divergences there is a gap between Hume's and Newton's ideologies of knowledge, i.e. between the frames within which they see the significance and meaning of inquiry and how it derives inspiration and legitimacy. The dominant stance of Newtonians in Scotland and abroad was in concert with the dictum in the General Scholium in the *Principia*: "to treat of God from phenomena is certainly a part of natural philosophy" (Newton 2004, 92). In this spirit many of them were working on fulfilling Newton's prophecy at the end of the *Opticks*, that following his method of analysis and synthesis "the Bounds of Moral Philosophy will also be enlarged." This enlargement should proceed through the perfection of natural philosophy, which consists in its increasing contribution to our knowledge of the attributes and intentions of God: "For so far as we can know by natural philosophy what is the first cause, what power he has over us, and what benefits we receive from him, so far our duty towards him, as well as that towards one another, will appear to us by the light of nature" (Newton 2004, 140).

Hume did not share the high hopes of Scottish "providential naturalism" for epistemic benefits concerning transcendence (for a summary see Broadie 2004, 41–43). The articulation of his stance in this respect is a leading theme of the *Enquiry*, and it might be labelled as *methodological agnosticism* that undermines the epistemic authority of moral (EHU Section 8/2) and natural inquiries (EHU 11) into the matters of transcendence, and of scriptural (or

13 This is an intended allusion to Kant's "transcendental deduction" of the categories. On some of the affinities and divergences of Hume's and Kant's project see Demeter 2016, ch. 3.

Biblical) inquiries (EHU 10) into matters of nature. The common lesson that Hume can offer to them is that they cannot be sources of epistemic value. This agnosticism is typically not expressed in explicit methodological commitments, but it is exhibited throughout various contexts. Hume comes close to a clear formulation of this stance when he says that “[w]e can never be allowed to mount up from the universe, the effect, to Jupiter, the cause; and then descend downwards, to infer any new effect from that cause” (EHU 11.14). It is easy to read this as a ban, contrary to Newton’s advice in Query 31 of the *Opticks* on extending the method of analysis and synthesis so as to reach conclusions concerning transcendent matters. As Hume considers this method the only appropriate method of reasoning about empirical matters, he thereby renders the properties of transcendent existence inaccessible by experimental reasoning (for more details see Demeter 2016: ch. 5 and 179–182).

For Hume, the “only immediate utility of all sciences, is to teach us, how to controul and regulate future events by their causes” (EHU 7.29). This is common to moral and natural philosophy, and this is how we can expect improvement in society from the cultivation of the “accurate”, “abstruse”, “profound” and “subtile” forms of philosophy (EHU 1.4–5, 9). But beyond this instrumental utility (and the intellectual pleasure we can derive from pursuing philosophical enterprises (see e.g. EHU 1.10, T 2.3.10.4)), we are in vain to extend the boundaries of science to transcendent matters. Even if God is the author of the Book of Nature, he did not equip us with the necessary tools for reading it. Hume thus rejects Newton’s vision concerning the study of God through nature and the enlargement of moral philosophy through our improved knowledge of the first cause.

## 5. Conclusion

In this paper I could take only the first steps towards an explication of Hume’s epistemic ideal by outlining its critical edge directed against then-influential alternatives. The clarification of central epistemic concepts and the epistemic burden they can carry, as well as the prospects and limits of the epistemic genres organized by them, must be left to other occasions. Still, the picture drawn here, despite being only partial, can still reveal some central features of Hume’s epistemic ideal. Close relatives of these features are easily discernible in the self-image of modern science that emerged gradually as a result of putting forward, and putting into use, alternative conceptions and practices of knowledge production.

These features include Hume’s empiricism which is moderately optimistic about advancing knowledge claims with respect to unobserved structures and processes, and completely pessimistic in those respects where no empirical analogies can be exploited for analytical and inferential purposes. Although empirical inquiry can only aspire to a degree of certainty

that falls short of being demonstrative, it can still be certain enough to be unanimously reliable and instrumentally useful. The Humean ideal does not favour mathematization: although mathematical representation and analysis of phenomena is not against Hume's taste, mathematics cannot be the guide or model of inquiry and discovery into matters of fact. Mathematics is a field of knowledge and relies on a form of reasoning that is categorically distinct from empirical reasoning (see Demeter 2019b, Slavov 2017), so it develops in complete isolation from the empirical sciences. Against the background of Hume's epistemic ideal, the thorough mathematization of natural philosophy in the following centuries could have testified only what Wigner (1960) famously termed "the unreasonable effectiveness of mathematics in the natural sciences".

Yet, the main lesson Hume offers is ideological and has been incorporated into the self-image of modern science: insofar as we study nature in accordance with sober methodological rules (to which Newton and his followers assents as well), we cannot hope for any knowledge of transcendent matters. If we are to stick to the, by Newton's standards, best method of reasoning that the "nature of things admits of" – namely induction based on experience, observation and analogy – then we have to give up the Newtonian ideology of natural philosophy: its religious frame of significance is to be replaced by an entirely secular one – that of Humean instrumentality.

## References

- CSM – *The Philosophical Writings of Descartes*, Vols. 1 and 2, translated by John Cottingham, Robert Stoothof, and Dugald Murdoch, Cambridge: Cambridge University Press, 1985.
- EHU – *An Enquiry concerning Human Understanding*, ed. Tom L. Beauchamp, Oxford: Oxford University Press, 2000.
- T – *A Treatise of Human Nature*, 2 vols., eds. David Fate Norton and Mary J. Norton, Oxford: Oxford University Press, 2007.
- Barfoot, Michael, 1990. "Hume and the Culture of Science" in M.A. Stewart (ed.), *Studies in the Philosophy of Scottish Enlightenment*, Oxford: Oxford University Press.
- Broadie, Alexander, 2004. "Reid in Context" in T. Cueno and R. van Woudenberg (eds.), *The Cambridge Companion to Thomas Reid*, Cambridge: Cambridge University Press, 31–52.
- Buckle, Stephen, 2001. *Hume's Enlightenment Tract*, Oxford: Clarendon Press.
- Cohen, I.B., 1980, *The Newtonian Revolution*, Cambridge: Cambridge University Press.

- Coventry, Angela, 2006. *Hume's Theory of Causation: A Quasi-Realist Interpretation*, London: Continuum.
- Demeter, Tamás, 2016. *David Hume and the Culture of Scottish Newtonianism: Methodology and Ideology in Enlightenment Inquiry*, Leiden: Brill.
- Demeter, Tamás, 2019a. "Hume's Science of Mind and Newtonianism" in Eric Schliesser and Christopher Smeenk (eds.), *The Oxford Handbook of Newton*, <https://doi.org/10.1093/oxfordhb/9780199930418.013.19>
- Demeter, Tamás, 2019b. "Hume on the Social Construction of Mathematical Knowledge", *Synthese* 196/9, 3615–3631.
- Demeter, Tamás, 2020. "The Science in Hume's Science of Man", *Journal of Scottish Philosophy*, 18/3, 257–271.
- Demeter, Tamás, 2021. "Fodor's Guide to the Humean Mind", *Synthese*, <https://doi.org/10.1007/s11229-021-03028-4>
- Demeter, Tamás, Benedek Láng and Dániel Schmal, 2016. "Scientia" in Marco Sgarbi (ed.), *Encyclopedia of Renaissance Philosophy*, Cham: Springer, [https://link.springer.com/referenceworkentry/10.1007%2F978-3-319-02848-4\\_266-1](https://link.springer.com/referenceworkentry/10.1007%2F978-3-319-02848-4_266-1)
- Garber, Daniel, 1992. *Descartes' Metaphysical Physics*, Chicago: University of Chicago Press.
- Garber, Daniel, 2001. *Descartes Embodied*, Cambridge: Cambridge University Press.
- Gaukroger, Stephen, 2014. "Empiricism as a Development of Experimental Natural Philosophy" in Zvi Biener and Eric Schliesser (eds.), *Newton and Empiricism*, New York: Oxford University Press.
- Guicciardini, Nicoló, 2009. *Isaac Newton on Mathematical Certainty and Method*, Cambridge, Mass.: MIT Press.
- Harris, James A., 2015. *Hume: An Intellectual Biography*, Cambridge: Cambridge University Press.
- Henry, John, 2011. "Gravity and *De Gravitatione*: The Development of Newton's Ideas on Action at a Distance", *Studies in History and Philosophy of Science* 42/1, 11–27.
- Jardine, Lisa, 1974. *Francis Bacon and the Art of Discourse*, Cambridge: Cambridge University Press.
- Landy, David 2018. *Hume's Science of Human Nature: Scientific Realism, Reason, and Substantial Explanation*. London: Routledge.
- Newton, Isaac, 1672. "A Letter of Mr. Isaac Newton... Containing His New Theory of Light and Colours," *Philosophical Transactions of the Royal Society* 6, No. 80, 3075–87.

- Newton, Isaac, 2004. *Philosophical Writings*, Cambridge: Cambridge University Press.
- Norton, David Fate and Norton, Martha, 2007. "Historical Account of a *Treatise of Human Nature* from its Beginnings to the Time of Hume's Death." In Hume, *T*, vol. 2, Oxford: Clarendon Press, 433–588.
- Ott, Walter, 2009. *Causation and Laws of Nature in Early Modern Philosophy*, Oxford: Oxford University Press.
- Pasnau, Robert, 2013. "Epistemology Idealized", *Mind* 122/488, 987–1022.
- Pasnau, Robert, 2017. *After Certainty: A History of Our Epistemic Ideals and Illusions*, Oxford: Oxford University Press.
- Perinetti, Dario, 2018. "Hume at La Flèche: Skepticism and the French Connection", *Journal of the History of Philosophy* 56/1, 45–74.
- Qu, Hsueh M., 2020. *Hume's Epistemological Evolution*, Oxford: Oxford University Press.
- Reid, Thomas, 1872. *The Works of Thomas Reid*, 2 vols. 7<sup>th</sup> ed. Edinburgh: Maclachlan and Stewart.
- Salmieri, Gregory, 2014. "Aristotelian Episteme and the Relation between Knowledge and Understanding", *Metascience* 23/1, 1–9.
- Schliesser, Eric, 2011, "Newton's Challenge to Philosophy: A Programmatic Essay," *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 1/1, 101–28.
- Schliesser, Eric and Demeter, Tamás, 2020. "Hume's Newtonianism and Anti-Newtonianism", *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/hume-newton/>
- Sgarbi, Marco, 2013. *The Aristotelian Tradition and the Rise of British Empiricism: Logic and Epistemology in the British Isles (1570–1689)*, Cham: Springer.
- Shapiro, Alan E., 1993. *Fits, Passions, and Paroxysms: Physics, Method, and Chemistry and Newton's Theories of Colored Bodies and Fits of Easy Reflection*, Cambridge: Cambridge University Press.
- Shepherd, Christine M., 1982. "Newtonianism in Scottish Universities in the Seventeenth Century", in R.H. Campbell and A. Skinner (eds.), *The Origins and Nature of the Scottish Enlightenment*, Edinburgh: John Donald.
- Slavov, Matias, 2016. "Newtonian and Non-Newtonian Elements in Hume", *Journal of Scottish Philosophy* 14/3, 275–296.
- Slavov, Matias, 2017. "Hume's Fork and Mixed Mathematics", *Archiv für Geschichte der Philosophie* 99/1, 102–119.



- Smith, George E., 2002. "The Methodology of the *Principia*" in I.B. Cohen and George E. Smith (eds.), *The Cambridge Companion to Newton*, Cambridge: Cambridge University Press, 138–73.
- Smith, Kurt, 2010. *Matter Matters: Metaphysics and Methodology in the Early Modern Period*, Oxford: Oxford University Press.
- Smith, Robin, 2017. "Aristotle's Logic", *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/aristotle-logic/>
- Stewart M.A., 2005, "Hume's Intellectual Development" in Marina Frasca-Spada and P.J.E. Kail (eds.), *Impressions of Hume*, Oxford: Oxford University Press.
- Turnbull, H. W., ed. 1959. *The correspondence of Isaac Newton I. 1661–1675*. Cambridge: Cambridge University Press.
- Turnbull, H. W., ed. 1960. *The correspondence of Isaac Newton II. 1676–1687*. Cambridge: Cambridge University Press.
- Wigner, Eugene, 1960, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences", *Commun. Pure Appl. Math.* 13/1, 1–14.
- Wilson, David B., 2009. *Seeking Nature's Logic: Natural Philosophy in the Scottish Enlightenment*, University Park: Penn State University Press.
- Wright, John P., 2009. *Hume's A Treatise of Human Nature: An Introduction*, Cambridge: Cambridge University Press.
- Zemplén, Gábor Á. and Demeter, Tamás, 2010. "Being Charitable to Scientific Controversies: On the Demonstrativity of Newton's *Experimentum Crucis*", *Monist* 93/4, 640–656.



## HUME AND RELIABILISM

**Abstract:** *Hume's epistemological legacy is often perceived as a predominantly negative sceptical one. His infamous problem of induction continues to perplex philosophers to this day, and many of his sceptical worries maintain their interest in contemporary eyes (e.g. with regard to reason, the senses, substance, causation). Yet Hume's positive epistemological contributions also hold significance for philosophy in this day and age. In this paper, I aim to situate Hume's epistemology in a more contemporary context, particularly with regard to the theme of reliabilism that runs throughout this epistemology. This will take the shape of examining correspondences and contrasts between Hume's epistemologies in the Treatise and Enquiry and reliabilism, as well as an examination of how Hume's framework might handle some major challenges for reliabilist epistemologies. In particular, I argue that that while Hume is tempted to an epistemology that is intimately tied to truth in the Treatise, he backs away when confronted with the excesses of scepticism in the conclusion of Book 1, and winds up with an epistemology most similar to the contemporary epistemological frameworks of dogmatism and phenomenal conservatism. Yet, largely because of his reliance on the passions (a respect in which he diverges from these two contemporary frameworks), the epistemology of the Treatise remains crucially dissociated from truth. Meanwhile, in the first Enquiry, he proceeds to develop a two-tiered epistemological framework that first accords all our justification with default authority, and then founds all-things-considered epistemic justification on our evidence for the reliability of our faculties. The first tier most resembles the contemporary epistemological framework of conservatism, while the second tier most closely resembles approved-list reliabilism. In this, a clear reliabilist thread runs through the epistemology of the Enquiry. I will also argue that although Hume did not appear to fully appreciate one of the most significant challenges for reliabilism—that is, the generality problem—his philosophical framework nevertheless contains the beginnings of a response to it.*

**Keywords:** *Hume, scepticism, externalism, internalism, reliabilism*

### 1. Introduction

Hume's epistemological legacy is often perceived as a predominantly negative sceptical one. His infamous problem of induction continues to perplex philosophers to this day, and many of his sceptical worries maintain their interest in contemporary eyes (e.g. with regard to reason, the senses, substance, causation). Yet Hume's positive epistemological contributions also hold significance for philosophy in this day and age. In this paper, I aim to situate Hume's epistemology in a more contemporary context,

particularly with regard to the theme of reliabilism that runs throughout this epistemology. This will take the shape of examining correspondences and contrasts between Hume's epistemologies in the *Treatise* and *Enquiry* and reliabilist frameworks, as well as an examination of how Hume's system might handle some major challenges for reliabilist epistemologies. This paper thus builds on my previous monograph investigating the epistemological differences between the *Treatise* and the *Enquiry*.<sup>1</sup> While the core of my position has not substantively changed, this paper, unlike the monograph, explores the differences between these works through the particular lenses of reliabilism. It also develops my views in various respects, notably in more deeply exploring Hume's correspondences with contemporary epistemological frameworks, and formulating a more thorough response on Hume's behalf to the generality problem. Notably, this paper also represents a shift in my own views on Hume and externalism: unlike in my earlier work, this paper sees the *Enquiry* as having externalist tendencies, suitably qualified.<sup>2</sup>

Intuitively, epistemological justification must be, in some way, shape or form, related to truth. What is the point of pursuing epistemic justification if doing so does not make us more likely to believe true things? Reliabilism looks to draw as straight as possible a line between these two notions.<sup>3</sup> There is a tremendous variety of such accounts, but very broadly, process reliabilism takes a belief to be justified if it proceeds from a reliable or truth-conducive process.

Process reliabilism is typically taken to be externalist in nature—after all, what matters isn't one's evidence for a belief, but whether the belief proceeds from a process that is in fact reliable.<sup>4</sup> However, this may be disputed, as we will see. For one, a number of reliabilist accounts have adopted internalist-friendly elements. And some internalist accounts have adopted the mantle of reliabilism. In this paper, I use 'reliabilism' to encompass internalist accounts that tie epistemological justification to truth-conduciveness in an intimate way. Even if one wishes to maintain that process reliabilism is a distinctively externalist account, it cannot be denied that it has a core insight that can be adopted by internalist accounts, which is that epistemological justification closely concerns in some respect the truth-conduciveness of doxastic processes.

---

1 (Qu, 2020).

2 (Qu, 2018b) and (Qu, 2020).

3 Reliabilism has gained a great deal of prominence in contemporary epistemology. The classic statement of reliabilism about justification is (Goldman, 1979). Important early statements of reliabilism about knowledge are (Goldman, 1967) and (Dretske, 1981).

4 How to characterise the internalist/externalist distinction is itself a matter of some dispute, but for the purposes of this paper a sensible way to draw the distinction might be as follows: an epistemology is internalist iff a person either has access or potential access to the basis for their epistemic justification. This form of internalism is referred to as 'accessibilism' in (Feldman & Conee, 2001). (Beddor and Goldman, 2015) argue that since 'being reliably caused' is neither a mental state nor directly accessible to the believer, process reliabilism must be an externalist theory.

Hume agrees with the process reliabilist about the epistemic primacy of doxastic processes, as opposed to, say, evidence or belief. In the moral case, Hume argues that strictly speaking, it is not our actions that are the basis for moral evaluation, but rather the underlying characters that produce them (THN 2.3.2.6, THN 3.3.1.4; SBN 411, 575).<sup>5</sup> For Hume, moral evaluations are based on durable character traits. The intuition here seems to be something like as follows. A person cannot be praised or blamed at time *t* for a quality that they do not possess or instantiate at time *t*. Since actions are transient, we cannot blame someone for an action they have committed in the past. However, moral evaluations are more persistent than actions, and thus the latter cannot be the basis for the former. Instead, the durable basis upon which we are to be morally evaluated is the underlying stable character traits or dispositions that give rise to these actions. Since we cannot directly observe these character traits, they must be inferred on the basis of a person's actions (EHU 8.31; SBN 99).<sup>6</sup>

Hume's epistemological framework is unfortunately much less explicit than his moral one.<sup>7</sup> But there seems little reason not to apply the same line of thought here. Beliefs are lively ideas (THN 1.3.7.6; SBN 97); being occurrent, they are, like actions, transient.<sup>8</sup> This lack of durability precludes them from being the basis of epistemic evaluation.<sup>9</sup> The proper basis of epistemic evaluation is instead the underlying stable dispositions or processes that produce these beliefs—reason, custom, the imagination, and so forth.<sup>10</sup>

We have seen that Hume agrees with the process reliabilist that epistemic justification turns on belief-forming processes. But does he agree that their truth-conduciveness is the crucial normative property of these processes? I will argue that he is tempted to such a view in the *Treatise*, although he backs away when confronted with the excesses of scepticism in the conclusion of Book 1. However, in the first *Enquiry*, he proceeds to develop a two-tiered epistemological framework that founds epistemic justification on our evidence for the reliability of our faculties. I will also argue that although

5 In references to Hume's texts throughout the paper, 'THN' refers to the *Treatise of Human Nature*, 'EHU' to the *Enquiry Concerning Human Understanding*, and 'EPM' to the *Enquiry Concerning the Principles of Morals*. Arabic numerals refer to section and paragraph numbers (EHU and EPM), or to book, part, section, and paragraph numbers (THN). SBN numbers refer to numbers in the Selby-Bigge edition of these works.

6 (Qu, 2017, p.84).

7 Perhaps this is because, as (Loeb, 2011) and (Ainslie, 2015) argue, Hume's epistemology occurs primarily in service of his psychology. I agree that this is true for the *Treatise* (Qu, 2020, p.26).

8 (Loeb, 2002) argues that Hume is fundamentally committed to a dispositional account of beliefs, but see (Marusic, 2010) for a convincing criticism of this view.

9 (Greco, 2012, Ch.9) argues for a virtue-theoretic epistemology on very similar grounds.

10 (Qu, 2014b, pp.515–516). (Schmitt, 2014) attributes this view to Hume on other grounds (i.e. inheriting it from his predecessors such as the Cartesians and Locke), calling it an 'operations-based epistemology' (p.28).

he did not appear to fully appreciate one of the most significant challenges for reliabilism—that is, the generality problem—his philosophical framework nevertheless contains the beginnings of a response to it.

## 2. Truth-conduciveness in the *Treatise*

Over the years, various commentators have defended reliabilist readings of the *Treatise*. (Costa, 1981) was an early proponent, and (Beebe, 2006, p.73) briefly suggests reading Hume as an inductive reliabilist. However, it is safe to say that none have done more to advance the reliabilist reading of Hume than (Schmitt, 2014). Briefly, Schmitt argues that Hume's epistemological motivations are to maintain that demonstrative reasoning and probable reasoning have comparable, and laudable, epistemic status. Yet, the two belief-forming processes are intrinsically very different (p.22), and so their epistemological commonality must derive from their extrinsic features instead. Schmitt proposes that the relevant extrinsic feature is *reliability*—although they issue from very different psychological processes, these processes have in common their tendency to produce true beliefs.

Rather than particular snippets of text, the strength of Schmitt's framework rests more on the overarching narrative it imbues upon Book 1, to which I cannot do justice here. Still, some brief textual evidence is as follows. Schmitt argues persuasively that beliefs have a natural function to facilitate the avoidance of calamities (THN 1.3.10.2; SBN 118–9), and that they fulfil this natural function by delivering truth.<sup>11</sup> Correspondingly, the natural function of our belief-forming processes is to deliver true beliefs.<sup>12</sup> Moreover, Hume frequently associates 'just' and its converse with truth and falsity (THN 1.2.4.17, THN 1.4.1.1, THN 1.4.7.13, THN 2.1.3.7, THN 2.3.10.2, THN App.1; SBN 41–2, 180, 271–2, 282, 448–9, 623).<sup>13</sup> More can be said on this point, but at the very least, it seems that Hume recognises that epistemology should not be divorced from truth.

For the most part, I agree that Hume in the *Treatise* seeks an epistemology closely tied with truth. Nevertheless, I will briefly adduce some reasons, sourcing from the conclusion of Book 1, to think that when confronted with the excesses of scepticism, Hume becomes decidedly pessimistic about the prospects for such an epistemology, and correspondingly retreats.<sup>14</sup>

Hume faces an epistemological reckoning in the conclusion of Book 1. He frames his primary sceptical worry as a dissociation from truth:

Can I be sure, that in leaving all establish'd opinions I am following truth; and by what criterion shall I distinguish her, even if fortune

---

11 (Schmitt, 2014, p.108).

12 (Schmitt, 2014, p.112).

13 (Schmitt, 2014, Ch.4).

14 (Qu, 2020, pp.159–166).

shou'd at last guide me on her foot-steps? After the most accurate and exact of my reasonings, I can give no reason why I shou'd assent to it; and feel nothing but a *strong* propensity to consider objects *strongly* in that view, under which they appear to me.... Without this quality, by which the mind enlivens some ideas beyond others (which seemingly is so trivial, and so little founded on reason) we cou'd never assent to any argument, nor carry our view beyond those few objects, which are present to our senses. (THN 1.4.7.3; SBN 265)

The concern is essentially that there does not seem to be a reliable link between our ideas being lively and their being true. Hume lists habit, the memory, the senses, and the imagination as being hostages to fortune in this respect. Essentially, virtually all our belief-forming processes share this epistemically unfortunate dependence of vivacity, leaving us with little reason to think they are indeed truth-conducive.

One thing to note at this juncture is that Hume's framing of this issue here indicates that his epistemological concern is internalist rather than externalist in nature. His worry is not whether our belief-forming processes are *in fact* reliable. His worry is rather that we have no *reason* (and here he seems to mean an *introspectively available reason*—'I can give no reason') to think that they are reliable. While this is perhaps not decisive, it is at least suggestive.

How then does Hume resolve his sceptical problems? I lack the space to enter into a detailed analysis of THN 1.4.7 (SBN 263–7); I do so elsewhere, and will be brief here.<sup>15</sup> Having been left with no reason to think our belief-forming processes reliable, Hume is at a loss as to which of them to depend on. This culminates in the 'dangerous dilemma' (THN 1.4.7.6–7; SBN 267–8), which can be stated as follows: should we assent only to the trivial suggestions of the imagination, or should we instead adhere wholly to the dictates of the understanding? The former (the 'credulity horn') leads to 'errors, absurdities, and obscurities,' while the latter (the 'sceptical horn') leads to the utter annihilation of our beliefs, given the sceptical regress of THN 1.4.1 (SBN 180–7), which can only be avoided by the aforementioned trivial propensities. Solutionless, Hume stumbles into melancholy and delirium (THN 1.4.7.9; SBN 269), followed by a period of 'spleen and indolence' (THN 1.4.7.11; SBN 270). Near the end of this period, Hume presents what has, following Garrett, become known as the 'Title Principle':<sup>16</sup>

15 (Qu, 2014b), (Qu, 2020, Chs. 6–7).

16 While controversial, the centrality of the Title Principle to Hume's response to scepticism has many defenders in the literature. See for instance Garrett in (Garrett, 1997, pp.233–237), (Garrett, 2006), (Garrett, 2015, pp.27–237), and (Garrett, 2016); (Mounce, 1999, p.60); (Kail, 2007, p.70); (Allison, 2008, pp.323–330); (Meeker, 2013, pp.73–81); (Qu, 2014b) and (Qu, 2020); (Schafer, 2014); (Schmitt, 2014, pp.368–375); (Baxter, 2018, pp.388–389), and (Sasser, Forthcoming).

Where reason is lively, and mixes itself with some propensity, it ought to be assented to. Where it does not, it never can have any title to operate upon us. (THN 1.4.7.11; SBN 270)

Notably, the Title Principle navigates the two horns of the dangerous dilemma. The credulity horn is avoided, since we do not wholly assent to the belief-forming trivial propensities of the imagination, but allow for reason to play a central role. And the sceptical horn is avoided, since the iterative reasoning that engenders the doxastic apocalypse is devoid of any liveliness, and is shunned by our propensities. Thus, the Title Principle licenses Hume's return to philosophical reasoning, when conjoined with the return of his passions or curiosity and ambition (THN 1.4.7.12; SBN 270–1): such reasoning is lively and mixes with these two propensities.<sup>17</sup>

The Title Principle, in emphasising the role of the liveliness or vivacity of our beliefs, has affinities with certain brands of permissive foundationalism that tie justification to a certain phenomenal character. Most prominent among this family of views include Huemer's 'phenomenal conservatism', and Pryor's 'dogmatism'.<sup>18</sup> Phenomenal conservatism holds that epistemic justification is conferred by seemings or appearances—the fact that it appears as though P provides some *prima facie* justification for P. The notion of a seeming has distinct parallels with Hume's notion of liveliness, since to have a lively idea (that is, a belief) is for the idea to appear as true. And like vivacity (THN 1.3.7.7; SBN 628–9), seemings are phenomenal,<sup>19</sup> and unanalysable.<sup>20</sup> Just as a seeming carries epistemic justification for Huemer, so too does liveliness as per the Title Principle. Of course, there are differences. Notably, Huemer treats seemings as distinct mental states from beliefs, whereas Hume takes vivacity to be constitutive of belief.<sup>21</sup> But the core insight, that some mental states enjoy epistemic justification at least partly in virtue of their phenomenal force, seems to be shared between the Title Principle and phenomenal conservatism.

17 Thus, by 'propensities', Hume seems to mean the passions. See (Schafer, 2014, p.9) and (Qu, 2020, p.128), but refer to (Durland, 2011, pp.82–83) for disagreement.

18 See (Huemer, 2001), (Huemer, 2007), (Huemer, 2013) and (Pryor, 2000), (Pryor, 2004). Similar views are defended by (Tucker, 2010), (Brogaard, 2013), and (Chudnoff, 2013), among others.

19 Huemer describes appearances as having 'forcefulness' (Huemer, 2001, p.77) and 'assertive' character (Huemer, 2013, p.329), which suggests their having a phenomenal dimension. However, he denies that 'appears' takes a phenomenal use, although this seems driven by the implausibility of appearances always carrying a certain sensory quale (Huemer, 2013, p.330), which is consistent with appearances nevertheless having a non-sensory phenomenal feel. (Tolhurst, 1998, p.298) refers to the 'felt veridicality' of seemings, and the phenomenal dimension of seemings is asserted by (Tucker, 2010, p.530), (Werner, 2014, p.1765), and (McAllister, 2018, p.3082), to name a few.

20 (Huemer, 2013, pp.328–329).

21 (Huemer, 2007, p.31).



Meanwhile, dogmatism is the position that our perceptual beliefs are immediately *prima facie* justified by their phenomenology. As with seemings, the distinctive justification-conferring phenomenology is one that presents its contents as true.<sup>22</sup> Again, there are differences: dogmatism typically takes only perceptual beliefs to enjoy such justification, whereas the Title Principle extends its domain to all beliefs.<sup>23</sup> But again, there is a shared core insight between the two: certain phenomenologies can confer epistemic justification.<sup>24</sup>

Note that the Title Principle, like phenomenal conservatism and dogmatism, has a distinctly internalist character.<sup>25</sup> It founds epistemic justification on vivacity and the passions. The former component, insofar as it is an experiential component of our perceptions, is clearly introspectable. As for the latter, our passions have a distinctive qualitative character, insofar as they are impressions of reflection (THN 1.1.2.1, THN 1.1.6.1, THN 2.1.1.1; SBN 7–8, 15–16, 275).<sup>26</sup> The factors that confer epistemic justification are ones that are mentally available to us.

At the same time, the Title Principle seems to veer away from the straight and narrow path of an epistemology founded on truth-conduciveness. For one, Hume does not justify the Title Principle on the basis that the relevant belief-forming process (of reason mixing with some propensity) is reliable; indeed, he does not offer any justification for it at all.<sup>27</sup>

Moreover, while we examined the contemporary significance of Hume's tying epistemic justification to liveliness via the Title Principle, it should not be forgotten that the Title Principle also insists on tying epistemic justification to our propensities or passions, a move which has a markedly less storied legacy in contemporary epistemology. Importantly, our propensities do

22 (Pryor, 2004, p.357).

23 (Teng, 2018) has recently objected to dogmatism on the basis that the justification-conferring phenomenology of perceptions is one that is psychologically shared with mere imaginings, and this feature of imaginings fails to confer justification. This psychological observation coheres well with Hume's framework, whereby the difference between merely imagined ideas and beliefs boils down to the degree of vivacity present in the perception.

24 The link between Pryor's dogmatism and the Title Principle is observed in (Garrett, 2015, p.244), (Qu, 2014a, p.28), and (Qu, 2020, p.222).

25 Hume would probably not be very friendly towards reliabilist twists on such views that relate seemings with truth-conduciveness. For Hume himself casts major doubt on whether vivacity can be said to meaningfully co-vary with truth (THN 1.4.7.3).

26 Some commentators have argued that the calm passions lack any qualitative character whatsoever, for instance (Stroud, 1977, p.164), (Shaw, 1989), and (Smith, 1994, p.112). I have argued in (Qu, 2018a) that this is incorrect—although they are less emotionally intense than the violent passions, Hume says that they have 'little emotion', and not 'no emotion' (THN 2.3.3.8); moreover he explicitly takes them to be prone to confusion with the determinations of reason precisely because 'their *sensations* are not evidently different' (*ibid.*, emphasis added).

27 (Durland, 2011, pp.83–84), (Millican, 2016, p.105), (Qu, 2020, pp.147–152).

not seem particularly truth-conducive. Bill's desire for his daughter to be successful might cause him to falsely believe that she is doing better than she is; Mary's ambition might cause her to falsely believe unreliable gossip that she is to be promoted. An epistemology that accords our propensities a central role seems to have little in common with reliabilism. Importantly, this abandonment of truth-conduciveness does not appear to be accidental. Consider the problematic argument of THN 1.4.1 (SBN 180–7) that engendered the sceptical horn of the dangerous dilemma. Hume never takes issue with the argument's cogency, but only argues that it cannot have any psychological effect (THN 1.4.1.7; SBN 183). And the same is true of his sceptical argument with regard to the senses in THN 1.4.2 (SBN 187–218):

Tis impossible upon any system to defend either our understanding or senses; and we but expose them farther when we endeavour to justify them in that manner. As the sceptical doubt arises naturally from a profound and intense reflection on those subjects, it always increases, the farther we carry our reflections, whether in opposition or conformity to it. Carelessness and in-attention alone can afford us any remedy. (THN 1.4.2.57; SBN 218)

It seems that Hume blunts the sceptical force of these arguments not by denying their truth, but by appealing to our propensities, which, notoriously, are epistemically unreliable.

This is borne out in THN 1.4.7 (SBN 180–7). Just prior to endorsing the Title Principle, Hume notes that philosophical reasoning does not offer 'any tolerable prospect of arriving... at truth and certainty' (THN 1.4.7.10; SBN 269–70). And at the conclusion of THN 1.4.7 (SBN 180–7), Hume resigns himself to a philosophical system that is dissociated from truth:

But were these hypotheses once remov'd, we might hope to establish a system or set of opinions, which if not true (for that, perhaps, is too much to be hop'd for) might at least be satisfactory to the human mind, and might stand the test of the most critical examination. (THN 1.4.7.14; SBN 272)

Here, Hume goes so far as to say that the establishment of a true system is 'too much to be hoped for'—all that we can aspire to is that such a system stands up to 'the most critical examination.'<sup>28</sup>

This is of course not to say that Hume's reasonings cannot be true. But the crucial point is that his epistemology dissociates epistemic justification from truth in grounding it on factors that are not truth-tracking. To sum up,

28 Hume does hope to 'contribute a little to the advancement of knowledge', but this contribution does not lie in the discovery of truth *per se*, but only in 'giving in some particulars a different turn to the speculations of philosophers, and pointing out them more distinctly those subjects, where alone they can expect assurance and conviction' (THN 1.4.7.14; SBN 273).

in the *Treatise*, Hume reluctantly abandons the prospects of an epistemology founded on reliability. He does not abandon an externalist epistemology, but only because it did not appear to be on the cards to begin with.

### 3. Reliabilism in the Enquiry

There is, I venture, a stronger and more sustained theme of reliabilism that runs through the first *Enquiry*. First, consider Section 5.<sup>29</sup> Having presented his famous argument regarding induction in Section 4, he begins Section 5 by seeking to reassure readers who might be worried about its sceptical implications:

Nor need we fear, that this philosophy, while it endeavours to limit our enquiries to common life, should ever undermine the reasonings of common life, and carry its doubts so far as to destroy all action, as well as speculation. Nature will always maintain her rights, and prevail in the end over any abstract reasoning whatsoever. Though we should conclude, for instance, as in the foregoing section, that, in all reasonings from experience, there is a step taken by the mind, which is not supported by any argument or process of the understanding; there is no danger, that these reasonings, on which almost all knowledge depends, will ever be affected by such a discovery. If the mind be not engaged by argument to make this step, it must be induced by some other principle of equal weight and authority; and that principle will preserve its influence as long as human nature remains the same. What that principle is, may well be worth the pains of enquiry. (EHU 5.2; SBN 41–2)

This passage reveals two things. First, the argument of Section 4 is not a merely psychological one, but has an epistemological dimension.<sup>30</sup> Hume here is addressing the concerns of readers who are worried about the scope of the scepticism of Section 4: he notes that his brand of academical philosophy will not ‘carry its doubts so far as to destroy all action, as well as speculation’—which is to say, it does carry doubts in the first place. Second, having denied induction any justification sourcing from reason in Section 4, Hume looks to address any profound sceptical implications by offering an alternative source of justification: ‘If the mind be not engaged by argument to

29 I also discuss these issues in (Qu, 2014a) and (Qu, 2020, Ch.4).

30 For instance, ‘descriptivist’ accounts of THN 1.3.6 that see it as merely making a psychological point can be found in (Broughton, 1983), (Garrett, 1997, pp.91–95), (Garrett, 1998), (Owen, 1999), (Noonan, 1999), and (Allison, 2008, Chapter 5). Normative interpretations include (Winkler, 1999), (Loeb, 2006), (Loeb, 2008), (Millican, 1995), (Millican, 2002), (Qu, 2014a), and (Dimech, Forthcoming). I have suggested that the *Treatise*’s version is descriptive while the *Enquiry*’s is normative in (Qu, 2020, Chs.3–4), but I bracket this issue for the purposes of this paper.

make this step, it must be induced by some other principle of equal weight and authority'. 'Authority' in particular certainly seems to carry significant normative import—since our inductive practices are not produced by reason, which would offer them a certain form of justification, Hume will argue that they are produced by another belief-forming process that carries as much justification (albeit of a different form) as reason.

This belief-forming process is 'Custom, or Habit' (EHU 5.5; SBN 43). On what basis does custom carry this equal authority to reason? Hume expounds on this at some length at the close of Section 5:

Here, then, is a kind of pre-established harmony between the course of nature and the succession of our ideas; and though the powers and forces, by which the former is governed, be wholly unknown to us; yet our thoughts and conceptions have still, we find, gone on in the same train with the other works of nature. Custom is that principle, by which this correspondence has been effected; so necessary to the subsistence of our species, and the regulation of our conduct, in every circumstance and occurrence of human life. (EHU 5.21; SBN 54–5)

In making an admittedly sardonic reference to the 'pre-established harmony' that obtains between our inductive beliefs and the progression of nature via the mechanism of custom, Hume is explicitly referencing the truth-conducive nature of this subfaculty: the beliefs produced by custom tend to be true, and this is why it is a principle of 'equal weight and authority' to reason. Indeed, if anything, Hume goes on to claim that custom is superior to reason in this respect:

...this operation of the mind, by which we infer like effects from like causes, and *vice versa*, is so essential to the subsistence of all human creatures, it is not probable, that it could be trusted to the fallacious deductions of our reason, which is slow in its operations; appears not, in any degree, during the first years of infancy; and at best is, in every age and period of human life, extremely liable to error and mistake. It is more conformable to the ordinary wisdom of nature to secure so necessary an act of the mind, by some instinct or mechanical tendency, which may be infallible in its operations, may discover itself at the first appearance of life and thought, and may be independent of all the laboured deductions of the understanding. (EHU 5.22; SBN 55)

Reason is 'fallacious' and 'extremely liable to error and mistake'; custom, in having a correspondence with the course of nature, demonstrates a marked superiority in this regard.

(Dauer, 1980, p.363) and (Kail, 2016, p.152) take the conclusion of Section 5 to indicate that Hume endorses a reliabilist account in the *Enquiry*. Meanwhile, (Loeb, 2008, p.117) takes it to establish an externalist epistemology, although he demurs about the particular form this externalist epistemology takes. While I take it to be clear that these passages endorse

custom on the basis of its truth-conduciveness, it is underdetermined whether Hume commits himself to an externalist or internalist epistemology here. It is true that he claims that ‘we are ignorant of those powers and forces’ that determine the regular course of nature (EHU 5.22; SBN 55). But we are *not* ignorant of the fact that custom has a reliable track record: he emphasises that we ‘we find’ this correspondence between our inductive beliefs and nature to have obtained (EHU 5.21; SBN 54), which is suggestive that what justifies our reliance on custom is not so much that it is in fact truth-conducive, but rather that we find it to be such. To discern the epistemological framework that grounds this endorsement, we will now turn to Section 12.<sup>31</sup>

Hume’s goal in this section is to determine what form of scepticism he should endorse. To this end, he dismisses a number of options, before finally endorsing his own mitigated scepticism. First, he dismisses *antecedent scepticism*, which doubts not only our beliefs, but also our faculties prior to any investigation as to their reliability:

There is a species of scepticism, *antecedent* to all study and philosophy, which is much inculcated by Des Cartes and others, as a sovereign preservative against error and precipitate judgment. It recommends an universal doubt, not only of all our former opinions and principles, but also of our very faculties; of whose veracity, say they, we must assure ourselves, by a chain of reasoning, deduced from some original principle, which cannot possibly be fallacious or deceitful. But neither is there any such original principle, which has a prerogative above others, that are self-evident and convincing: Or if there were, could we advance a step beyond it, but by the use of those very faculties, of which we are supposed to be already diffident. The Cartesian doubt, therefore, were it ever possible to be attained by any human creature (as it plainly is not) would be entirely incurable; and no reasoning could ever bring us to a state of assurance and conviction upon any subject. (EHU 12.3; SBN 149–50)

Hume raises a decisive objection to such *a priori* form of scepticism: without antecedently relying on our faculties, we could never hope to make any epistemic progress whatsoever. Thus, we are licensed to accord default or *prima facie* justification to our faculties.<sup>32</sup> We might call such justification *antecedent justification*.

This epistemological framework closely resembles a form of what is called ‘conservatism’ endorsed by Wright.<sup>33</sup> Wright suggests that we might be

31 See my (Qu, 2018b) and (Qu, 2020, Ch.9), which go into more textual depth, although this paper develops the contemporary significance of this text in more detail.

32 (Garrett, 2007, p.6), (Millican, 2012, p.59), (Qu, 2020, p.183).

33 (Wright, 2002), (Wright, 2004), (Wright, 2014). Wright distinguishes warrant or entitlement from justification, which is a distinction I will not make much of here. A view in this vicinity is endorsed by Burge, which argues that we have an unearned right

entitled to believe some foundational propositions on the basis of their being fundamental to our epistemological projects. These propositions, which he calls ‘cornerstones’ of a given region of thought, are characterised by being such that a lack of warrant for them entails a lack of warrant for any belief in this region.<sup>34</sup> Since doubting such cornerstones would be epistemically ruinous, the epistemically dominant strategy is to accept them—thus, we have warrant to do so, so long as we do not have sufficient reason to *disbelieve* the cornerstone in question.<sup>35</sup> Scepticism about these cornerstones are blocked because these we are non-evidentially entitled to them—the sceptic cannot cast doubt on our evidence for these propositions, because they require none.<sup>36</sup> These entitlements are typically taken to have an externalist dimension, insofar as our entitlement to cornerstones does not turn on any reasoning we might do, but rather hold in virtue of certain epistemic dependencies independent of our reasoning—thus, such entitlement is said to come ‘for free’.

In light of this, we might state Hume’s discussion of antecedent scepticism as follows. Consider the proposition ‘our cognitive faculties are reliable’. This proposition is a cornerstone—to doubt it would undermine any possibility of epistemic achievement whatsoever. And, prior to any investigation (the point at which antecedent scepticism enters the fray), we certainly lack any positive reason to think that this proposition is false. Thus, we are justified in holding this proposition, at least until sufficient positive reason emerges to doubt it. And, like with conservatism, this justification seems to have an externalist dimension.<sup>37</sup> For, if our reliance on our faculties were justified by a further belief that these faculties were crucial to our epistemological projects, the antecedent sceptic could well ask how we would know that *this* belief was justified. Without a prior justification of our faculties, we would be at a loss for an answer—trust in our faculties has to be bedrock.

---

to rely on certain belief-forming processes on the basis that such processes are reliably veridical; see (Burge, 1993), (Burge, 1996), (Burge, 2003). That said, Wright’s account seems to much better resemble Hume’s dismissal of antecedent scepticism.

34 (Wright, 2004, pp.167–168).

35 (Wright, 2004, p.189).

36 It has been objected that this can only offer pragmatic rather than epistemic reasons for accepting such cornerstones; for instance, see (Jenkins, 2007, p.27), (Pritchard, 2007, p.207), and (Coliva, 2015, p.68). I find persuasive Wright’s response, which is that pragmatic reasons are in general instrumental; with regard to cornerstones, these pragmatic reasons are in service of epistemic values, and so these pragmatic reasons are also epistemic in nature (Wright, 2014, p.239).

37 As mentioned in the introduction, previously in (Qu, 2018b) and (Qu, 2020, Ch.9), I have maintained that Hume’s epistemology in the *Enquiry* is internalist in character, but I have recently come to appreciate that it has externalist aspects as well. In particular, while my earlier work treated antecedent justification as internalist, I have come to think it is better characterised as externalist, for the reasons given above. We will later see another respect in which my views on externalism in the *Enquiry* have moved.

Importantly, antecedent justification is extremely weak, and has to be. It would be a terrible result if all our faculties were *ultima facie* justified—after all, we know from experience that some of our belief-forming processes are not entirely reliable. To flesh out the epistemological story, we need a story about *defeat*: the antecedent justification accorded to our faculties holds unless there is a relevant defeater. What would constitute such a defeater? Once the spectre of antecedent scepticism has been done away with, we now hold the tools with which to conduct a consequent investigation of our faculties. If this investigation finds that our faculties are not reliable (as judged by their own lights), then we will have a good reason to distrust our faculties, and their antecedent justification will be defeated. This leads us to what Hume calls *consequent scepticism*:

There is another species of scepticism, *consequent* to science and enquiry, when men are supposed to have discovered, either the absolute fallaciousness of their mental faculties, or their unfitness to reach any fixed determination in all those curious subjects of speculation, about which they are commonly employed. (EHU 12.5; SBN 150)

On the flip side, if this consequent investigation discerns that some of our faculties are indeed reliable, then their antecedent justification is affirmed. We might then say that these faculties enjoy *consequent justification*.

What constitutes sufficient reason to doubt the reliability of our faculties? Here, standards might vary, corresponding to different forms of scepticism. Hume distinguishes two forms of consequent scepticism: Pyrrhonian or excessive scepticism on the one hand, and mitigated scepticism on the other. I venture that these two forms of scepticism are to be distinguished on the basis of their different standards for defeat with respect to the antecedent justification accorded to our faculties. Hume does not offer an explicit statement of the methodology of Pyrrhonian scepticism, but if we examine its arguments, we will see that Pyrrhonian scepticism has a tendency to be too liberal with its standards for defeat: given a certain misstep from our faculties, the Pyrrhonian will take this to be decisive reason to dismiss it wholesale. Consider a number of arguments Hume produces on behalf of the Pyrrhonian.

The Pyrrhonian sceptic dismisses the reliability of our senses on the basis of the ‘trite topics’ (EHU 12.6; SBN 151) such as oars appearing bent in water, or double vision upon pressing one’s eye. But as Hume points out, this only demonstrates that the senses need to be corrected by reason, not that we cannot rely on them in any context: ‘These sceptical topics, indeed, are only sufficient to prove, that the senses alone are not implicitly to be depended on; but that we must correct their evidence by reason’ (*ibid.*).

The Pyrrhonian sceptic also doubts the veracity of our senses on the basis of scepticism arising from the basis of the doctrine of double existence: our instincts tell us that our perceptions of sensation are external objects (EHU 12.8; SBN 151–2), but reason contradicts this claim (EHU

12.9; SBN 152), leading philosophers to postulate that these perceptions are caused by external objects (EHU 12.11; SBN 152–3). But this is a thesis that is beyond the scope of our faculties, being a matter of fact that cannot be justified by experience (EHU 12.12; SBN 153). The doctrine of double existence also gives rise to the theory of primary and secondary qualities, which leads to problems with regard to abstraction (EHU 12.15; SBN 154–5). Yet we might respond: while these rarefied philosophical theses might be problematic, why should this cast any doubt on the first-order deliverances of our senses?

The Pyrrhonian sceptic also doubts all abstract reasoning merely on the basis of paradoxes regarding infinite divisibility (EHU 12.18; SBN 156–7). Granted that we might need to avoid abstruse reasoning about infinite divisibility, but surely the Pyrrhonian overreacts: can we not be confident of simple arithmetical calculations, for instance?

The Pyrrhonian sceptic, via the ‘popular’ argument, doubts all probable reasoning on the basis that there is a significant degree of variation amongst inductive judgments (EHU 12.21; SBN 158–9). But this is clearly an insufficient basis on which to doubt wholesale our probable reasoning: what about the many more instances whereby people agree, and are correct, in their probable judgments?

The Pyrrhonian sceptic, via the ‘philosophical’ argument, argues for the unreliability of custom, because it is an instinct, and ‘other instincts’ have proven ‘fallacious and deceitful’ (EHU 12.22; SBN 159). But this seems egregious: why should the fallaciousness of unrelated instincts bear on the reliability of custom? Indeed, as we saw at the close of Section 5, Hume thinks that we do have significant positive evidence for custom being truth-conducive.

Pyrrhonian scepticism is thus unsatisfactory.<sup>38</sup> Hume proposes in its place mitigated scepticism, which advocates the limitation of the scope of our enquiries:<sup>39</sup>

Another species of *mitigated* scepticism, which may be of advantage to mankind, and which may be the natural result of the Pyrrhonian doubts and scruples, is the limitation of our enquiries to such subjects as are best adapted to the narrow capacity of human understanding. (EHU 12.25; SBN 162)

---

38 Hume explicitly adduces pragmatic reasons to dismiss Pyrrhonian scepticism: namely its unsustainability, and its dangerousness were it sustainable (EHU 12.23; SBN 159–60). But these do not offer any *epistemic* reason for dismissing Pyrrhonian scepticism. Although the epistemic reasons are implicit, I think they are no less present.

39 Another aspect of mitigated scepticism is the general adoption of a degree of epistemic humility in lowering our credences across the board, since we are naturally disposed to overconfidence (EHU 12.24; SBN 161–2), which is sensible advice, but perhaps less systematically interesting for our purposes here.



In this, we can see mitigated scepticism as deriving naturally from the flaws of Pyrrhonian scepticism—Hume notes that such scepticism is ‘the result of... Pyrrhonism, or excessive scepticism, when its undistinguished doubts are, in some measure, corrected by common sense and reflection’ (EHU 12.24; SBN 161). Pyrrhonian scepticism, in taking any small imperfection to count as a relevant defeater for the reliability of that faculty, overgeneralises sceptical doubt on the basis of relatively thin evidence of the unreliability of our senses. Meanwhile, mitigated scepticism is more judicious in this respect. Hume’s exhortation to remain exactly within the boundaries of ‘subjects as are best adapted’ to our faculties chides the Pyrrhonian for its trigger-happy approach to consequent scepticism. In doubting wholesale the reliability of our faculties on such a meagre basis, the Pyrrhonian decidedly *understeps* the scope of our faculties. Particularly egregious are the cases whereby the Pyrrhonian doubts a faculty on the basis of enquiries that *overstep* the boundaries of human knowledge, such as scepticism arising from the doctrine of double existence, or scepticism regarding infinite divisibility. For these surely offer no reason to doubt our faculties with respect to more prosaic matters. In response to these two flaws of the Pyrrhonian, mitigated scepticism proposes a ‘Goldilocks’ solution: we should not overstep or understep the boundaries of our faculties, but remain exactly within their scope. Thus, mitigated scepticism recognises that many of our belief-forming processes enjoy consequent justification; lacking sufficient reason to think them unreliable, their antecedent justification is not defeated.

The above points to consequent scepticism having a reliabilistic strain running through it. Key to a belief-forming operation enjoying consequent justification is our verifying its truth-conduciveness. If we find a faculty or operation to be reliable, then its antecedent justification is not defeated; but if we find it to be unreliable, then its antecedent justification is defeated. Importantly, on such a view, it seems that evidence regarding a faculty’s reliability is *necessary* for its enjoying consequent justification.

At the same time, consequent justification seems to have an internalistic aspect to it. Mitigated scepticism, as a brand of consequent scepticism, arises as a result of an investigation into the reliability of our faculties, as Hume earlier notes in EHU 12.5 (SBN 150–1), and again in EHU 12.26 (SBN 163):

This narrow limitation, indeed, of our enquiries, is, in every respect, so reasonable, that it suffices to make the slightest examination into the natural powers of the human mind, and to compare them with their objects, in order to recommend it to us. (EHU 12.26; SBN 163)

Correspondingly, the consequent justification granted by mitigated scepticism likewise arises as a result of such enquiry, and thus is mentally available to us. What is crucial for epistemic justification is not so much whether our faculties are actually reliable, but whether we find them to be such.

I have previously argued that Hume's account of consequent justification resembles an epistemological framework called 'internalist reliabilism' (Steup, 2004).<sup>40</sup> Briefly, Steup argues that our perceptual experiences are sources of justification insofar as we have memory impressions of both the reliability of our senses and of our memory (p.408). While putting reliability at the forefront, this account is internalist insofar as justification derives not from the reliability of the underlying belief-forming processes that produce it *per se*, but rather our *evidence* of the reliability of these underlying processes, via our memory of their track record. Again, a prominent difference with Hume's framework is that, like many foundationalist accounts, Steup accords a special status to the senses and perceptual experiences, while Hume is concerned with our faculties more broadly. But there is a distinct resemblance between these two frameworks.

That having been said, a number of ostensibly externalist reliabilist accounts have adopted internalist-friendly elements. One such example is (Alston, 1988), which seeks to integrate evidentialism and reliabilism.<sup>41</sup> Notably for our purposes, while maintaining that epistemic justification sources from truth-conduciveness (p.278), Alston also maintains that the grounds for justification must be 'the *sort* of thing' that is available to the agent, even if not *in fact* thusly available (p.274).

In line with such accounts, I have come to think that another useful analogue for Hume's epistemology in the *Enquiry* is what has been called 'approved-list reliabilism'.<sup>42</sup> This view holds that beliefs are deemed to be justified if they issue from belief-forming processes that are designated as epistemically virtuous (they are 'on the approved list'), and these belief-forming processes are thusly designated if they are found to be reliable. This certainly seems a fair characterisation of Hume's framework in Section 12, as described above. Consequent investigation establishes which faculties are reliable and thus enjoy consequent justification. Those that do, go on the list, and beliefs that issue from them are correspondingly deemed justified.

---

40 (Qu, 2018b), (Qu, 2020, pp.223–224).

41 Other reliabilist views that incorporate internalist elements are as follows. (Tang, 2016) offers a development of Alston's account that applies to credences, which might be thought to fruitfully mirror the degrees of vivacity a Humean belief may possess. (Comesana, 2010) defends a view entitled 'evidentialist reliabilism', according to which a belief is justified if it is based on evidence, and the belief-forming process that produces it based on that evidence is reliable. Likewise, (Goldman, 2012) proposes an evidentialist framework that cashes out evidence partly in terms of reliability—this is a two-factor approach, whereby justification turns both on fit with experiential evidence and causation by a reliable experience-based process (p.146). In similar vein is (Henderson, Terry, & Potre, 2007). Rather than integrating reliabilism with evidentialism, (Beddor, Forthcoming) proposes an integration with a reasons-based framework, arguing that it makes for a superior account of defeat.

42 (Goldman, 1993), (Fricker, 2016). I think of it as 'Santa-reliabilism'—he's making a list, and checking it twice.

It is true that approved-list reliabilism is not in itself an internalist theory: what matters for epistemic justification is simply whether a particular belief-forming process is on the approved list or not. Yet such a theory still gives a crucial role to the investigation of the reliability of our faculties in determining what goes into this list, respecting the internalistic aspect of Hume's framework in the *Enquiry*.<sup>43</sup>

Still, there is a deeper and more intriguing connection between the two. 'Approved-list reliabilism', as Goldman defends it (although he does not use the term), is an exercise in 'descriptive epistemology', rather than a 'normative epistemology'.<sup>44</sup> A descriptive epistemology seeks to 'elucidate commonsense epistemic concepts and principles', while a normative epistemology looks to formulate 'a more adequate, sound, or systematic set of epistemic norms, in some way(s) transcending our naïve epistemic repertoire'.<sup>45</sup> In short, we might think of a descriptive epistemology as offering an account of justification-ascriptions, rather than justification proper.

However, for Hume, the gap between these two projects is narrower than it would be for most. In general, we might convert approved-list reliabilism into a normative theory by adding the normative claim that a belief is justified iff it is on the approved list of most normal speakers at the actual world. Hume's normative methodology suggests that this would be an addition that he would be comfortable with. When explaining his own positive epistemology, he explicitly states: 'philosophical decisions are nothing but the reflections of common life, methodised and corrected' (EHU 12.25; SBN 162). As I have argued elsewhere, Hume's normative project consists of the systematisation of our pre-philosophical normative judgments into a sound and cohesive system.<sup>46</sup> This is a direct result of Hume's account of normativity as essentially founded on sentiment. We deem an act to be moral because it derives from a virtue that excites sentiments of approbation; similarly with a wise judgment. Thus, to derive a true theory of morality or epistemology, Hume looks to systematise the deliverances of our normative sentiments. Indeed, Hume describes his methodology in the moral *Enquiry* as consisting of analysing 'that complication of mental qualities, which... we call Personal Merit' (EPM 1.10). He continues:

The only object of reasoning is to discover the circumstances on both sides, which are common to these qualities; to observe that particular in which the estimable qualities agree on the one hand, and the blameable on the other, and thence to reach the foundation of ethics, and find

---

43 This is another respect in which my views on externalism in Hume's *Enquiry* have moved, as alluded to earlier. While internalist reliabilism is an unapologetically internalist theory, approved-list reliabilism is an externalist one (albeit with internalist elements).

44 (Goldman, 1993, p.282).

45 (Goldman, 1993, pp.271–273).

46 The nuts and bolts of this systematising project are spelt out in (Qu, 2016).

those universal principles, from which all censure or approbation is ultimately derived. (EPM 1.10; SBN 174)

Hume's task is descriptive in Goldman's sense, but it is also normative. The two are intertwined: to describe and systematise our normative judgments is to produce a genuinely prescriptive epistemology. Just as he seeks to offer an account of virtue and vice in the moral domain via a study of our moral intuitions, he can be read as looking to provide an account of epistemic justification via a study of our epistemological intuitions.

#### 4. The Generality Problem

There are a number of worries at this juncture. First, it might be worried that there is a circularity involved in using our faculties to justify themselves. I have addressed this worry elsewhere,<sup>47</sup> but briefly, circularity is avoided because of the distinction between two types of justification: antecedent and consequent. Antecedent justification is leveraged in order to attain consequent justification, which avoids epistemic circularity.<sup>48</sup>

We have seen that there is a strain of reliabilist thought that runs through the *Enquiry* in particular. Yet, in this, Hume's framework in this later work correspondingly inherits the problems of reliabilism as well. I wish to conclude the paper by considering one such inherited worry—the generality problem.<sup>49</sup> The generality problem is often raised as an objection to reliabilism, although commentators have argued that its scope is much broader.<sup>50</sup> The problem at its heart is as follows. A given belief is the result of a process-token that can be typed in a large array of ways. My belief that there is a tree in front of me can be characterised as the result of perception, perception under such-and-such lighting, perception of tree-like objects, perception of natural objects, and so forth. Without a unique way of settling which process-type is relevant to a given epistemic evaluation, we will be unable to get a reliabilist theory off the ground.

47 (Qu, 2020, p.201).

48 Thus, this avoids accusations of illegitimate bootstrapping, which is problematic precisely because it embodies an epistemic circularity; for instance, see (Vogel, 2000). Hume does bootstrap, but it is benign. There are two further considerations to note in this regard. One is that Hume is not alone in this regard: (Cohen, 2002) has argued that bootstrapping afflicts all theories that take on a 'basic knowledge' structure (that is, theories that allow for a belief source to deliver knowledge prior to knowing that the source is reliable). Second, (Van Cleve, 2003) argues persuasively that bootstrapping cannot be avoided on pain of falling to excessive scepticism, and defends a form of it, indeed interpreting Reid as doing the same. This I take to be Hume's response as well. As Hume sees it, the choice is between antecedent scepticism and bootstrapping, and he chooses accordingly.

49 The classic statements of this problem are (Feldman, 1985) and (Conee & Feldman, 1998).

50 (Comesana, 2006), (Bishop, 2010), (Tolly, 2017).

This is a critical (arguably the most critical) problem for reliabilism and beyond, and it would be unrealistic to expect Hume to have an answer for it, centuries before either reliabilism or the generality problem were prominent. Indeed, there is little indication that he was aware of the problem at all; at most, he gestures at it when rejecting the policy of forbidding all abstruse reasoning partly on the basis that such a policy would be unprincipled in allowing some but not all principles of the imagination (THN 1.4.7.7; SBN 267–8).<sup>51</sup>

I have argued that nevertheless, Hume might have a putative response available: in this respect, epistemology should take its cue from psychology.<sup>52</sup> Hume proposes a ‘mental geography, or delineation of the distinct parts and powers of the mind’ (EHU 1.13; SBN 13); insofar as this endeavour is successful, we can use its results to determine the appropriate process-types. This is in line with some contemporary responses to the generality problem that appeal to psychological means of winnowing down process-types—process-types that do not play a causal role in the subject’s psychology should be dismissed.<sup>53</sup>

However, this does not seem to get us all the way to uniqueness, for it leaves open a range of psychological process-types.<sup>54</sup> For instance, take the belief that the mangosteen tree will bear fruit. Just according to Hume’s psychological framework, this belief could be described as issuing from custom, probable reasoning, reason, or the imagination (in the broad sense). All of these belief-forming processes carve at the psychological joints, according to Hume, and the generality problem remains.

Perhaps we can do a little better yet on Hume’s behalf. Recall from earlier that Hume’s epistemological project involved systematising our common-sensical epistemological intuitions. These intuitions will include intuitions about which belief-forming processes are epistemically relevant. For instance, the imagination (in the broad sense) will clearly seem too broad to be epistemically relevant to my belief that a mangosteen tree will bear fruit. We might fruitfully appeal to these intuitions to further triangulate the epistemically relevant belief-forming processes. This is akin to a solution offered by Schmitt, which argues that to solve the generality problem, it is not required that the reliabilist provide necessary and sufficient conditions for the relevance of types; rather, it suffices to simply provide some constraints on relevant factors.<sup>55</sup> This is because reliabilism can be supplemented by our

51 Although (Schmitt, 1992, p.140) sees this problem as weighing more heavily on Hume’s mind than I do.

52 (Qu, 2020, p.202).

53 (Alston, 1995), (Beebe, 2004), (Kampa, 2018), (Lyons, 2019).

54 (Conee & Feldman, 1998, pp.10–18). Beebe, Kampa, and Lyons offer sophisticated versions of the psychological approach that seek to deliver uniqueness, but these are quite distant from anything we can find in Hume.

55 (Schmitt, 1992, Ch.6), although see (Conee & Feldman, 1998, pp.18–20) for criticism.

common-sensical epistemological intuitions as to the relevance of processes.<sup>56</sup> In Hume's case, he has derived a systematised account of these intuitions, and he is licensed to make use of it in resolving the generalist problem without begging the question.

Importantly, by synthesising this *normative* carving of our belief-forming processes with Hume's *psychological* carving of the same, we can triangulate the relevant belief-forming processes for a given belief. This narrows down the field of options considerably more than merely psychological solutions. Is this cross-referencing approach sufficient to solve the generality problem? Perhaps, perhaps not. But as we have seen, it would be unreasonable to expect Hume to have a complete solution to this problem to hand. At the very least, he has the resources to make progress on such a solution. And can we ask for much more than that of Hume? In many ways, his thought remains relevant to some of the most cutting-edge developments in philosophy today.<sup>57</sup>

## References

- Ainslie, D. (2015). *Hume's True Scepticism*. Oxford: Oxford University Press.
- Allison, H. (2008). *Custom and Reason in Hume: A Kantian Reading of the First Book of the Treatise*. Oxford: Oxford University Press.
- Alston, W. (1988). An Internalist Externalism. *Synthese*, 74(3), 265–283.
- Alston, W. (1995). How to Think about Reliability. *Philosophical Topics*, 23(1), 1–29.
- Baxter, D. (2018). A Pyrrhonian Interpretation of Hume on Assent. In D. E. Machuca & B. Reed (Eds.), *Skepticism: From Antiquity to the Present* (pp. 380–394). New York: Bloomsbury.
- Beddor, R. (Forthcoming). Reasons for Reliabilism. In J. Brown & M. Simion (Eds.), *Reasons, Justification, and Defeat*: Oxford University Press.
- Beebe, J. (2004). The Generality Problem, Statistical Relevance and the Tri-Level Hypothesis. *Noûs*, 38(1), 177–195.
- Beebe, H. (2006). *Hume on Causation*. London: Routledge.
- Bishop, M. A. (2010). Why the Generality Problem is Everybody's Problem. *Philosophical Studies*, 151(2), 285–298.
- Brogaard, B. (2013). Phenomenal Seemings and Sensible Dogmatism. In C. Tucker (Ed.), *Seemings and Justification: New Essays on Dogmatism and Phenomenal Conservatism* (pp. 270–289). New York: Oxford University Press.

56 (Schmitt, 1992, pp.141–142).

57 Thanks to Frederick Schmitt, Bob Beddor, and Tang Weng Hong for very helpful and detailed comments on an earlier draft of the paper.

- Broughton, J. (1983). Hume's Skepticism about Causal Inferences. *Pacific Philosophical Quarterly*, 64, 3–18.
- Burge, T. (1993). Content Preservation. *Philosophical Review*, 102(4), 457–488.
- Burge, T. (1996). Our Entitlement to Self-Knowledge. *Proceedings of the Aristotelian Society*, 96(1), 91–116.
- Burge, T. (2003). Perceptual Entitlement. *Philosophy and Phenomenological Research*, 67(3), 503–548.
- Chudnoff, E. (2013). *Intuition*. Oxford: Oxford University Press.
- Cohen. (2002). Basic Knowledge and the Problem of Easy Knowledge. *Philosophy and Phenomenological Research*, 65(2), 309–329.
- Coliva, A. (2015). *Extended Rationality: A Hinge Epistemology*. London: Palgrave Macmillan.
- Comesana, J. (2006). A Well-Founded Solution to the Generality Problem. *Philosophical Studies*, 129(1), 27–47.
- Comesana, J. (2010). Evidentialist Reliabilism. *Noûs*, 44(4), 571–600.
- Conee, E., & Feldman, R. (1998). The Generality Problem for Reliabilism. *Philosophical Studies*, 89, 1–29.
- Costa, M. A. (1981). Hume and Justified Belief. *Canadian Journal of Philosophy*, 11, 219–228.
- Dauer, F. W. (1980). Hume's Skeptical Solution and the Causal Theory of Knowledge. *Philosophical Review*, 89, 357–378.
- Dimech, D. (Forthcoming). Quasi-Realism and Inductive Scepticism in Hume's Theory of Causation. *Australasian Journal of Philosophy*.
- Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge: MIT Press.
- Durland, K. (2011). Extreme Skepticism and Commitment in the *Treatise*. *Hume Studies*, 37(1), 65–98.
- Feldman, R. (1985). Reliability and Justification. *The Monist*, 68(2), 159–174.
- Feldman, R., & Conee, E. (2001). Internalism Defended. *American Philosophical Quarterly*, 38(1), 1–18.
- Fricker, E. (2016). Unreliable Testimony. In B. Mclaughlin & H. Kornblith (Eds.), *Goldman and His Critics* (pp. 88–123). Oxford: Wiley Blackwell.
- Garrett, D. (1997). *Cognition and Commitment in Hume's Philosophy*. New York: Oxford University Press.
- Garrett, D. (1998). Ideas, Reasons, and Scepticism: Response to my Critics. *Hume Studies*, 24(1), 171–194.

- Garrett, D. (2006). Hume's Conclusions in 'Conclusion of this Book'. In S. Traiger (Ed.), *The Blackwell Guide to Hume's Treatise* (pp. 151–176). Oxford: Wiley-Blackwell.
- Garrett, D. (2007). Reasons to act and believe: naturalism and rational justification in Hume's philosophical project. *Philosophical Studies*, 132(1), 1–16. doi:10.1007/s11098-006-9055-y
- Garrett, D. (2015). *Hume*. New York: Routledge.
- Garrett, D. (2016). Hume on Reason, Normativity, and the Title Principle. In P. Russell (Ed.), *The Oxford Handbook of Hume* (pp. 32–53). New York: Oxford University Press.
- Goldman, A. (1967). A Causal Theory of Knowing. *Journal of Philosophy*, 64(12), 357–372.
- Goldman, A. (1979). What is Justified Belief? In G. Pappas (Ed.), *Justification and Knowledge* (pp. 1–25). Dordrecht: Reidel.
- Goldman, A. (1993). Epistemic Folkways and Scientific Epistemology. *Philosophical Issues*, 3, 271–285.
- Goldman, A. (2012). Towards a Synthesis of Reliabilism and Evidentialism? In A. Goldman (Ed.), *Reliabilism and Contemporary Epistemology: Essays* (pp. 123–150). New York: Oxford University Press.
- Greco, J. (2012). *Achieving Knowledge*. Cambridge: Cambridge University Press.
- Henderson, D., Terry, H., & Potre, M. (2007). Transglobal Evidentialism-Reliabilism. *Acta Analytica*, 22, 281–300.
- Huemer, M. (2001). *Skepticism and the Veil of Perception*. Lanham, MD: Rowman and Littlefield.
- Huemer, M. (2007). Compassionate Phenomenal Conservatism. *Philosophy and Phenomenological Research*, 74(1), 30–55.
- Huemer, M. (2013). Phenomenal Conservatism Über Alles. In C. Tucker (Ed.), *Seemings and Justification: New Essays on Dogmatism and Phenomenal Conservatism* (pp. 328–350). Oxford: Oxford University Press.
- Hume, D. (1739–1740/2007). *A Treatise of Human Nature* (D. F. Norton & M. J. Norton Eds.). Oxford; New York: Oxford University Press.
- Hume, D. (1748/2007). *An Enquiry Concerning Human Understanding* (P. Millican Ed.). Oxford: Oxford University Press.
- Hume, D. (1751/1998). *An Enquiry Concerning the Principles of Morals* (T. L. Beauchamp Ed.). Oxford: Oxford University Press.
- Jenkins, C. (2007). Entitlement and Rationality. *Synthese*, 157(1), 25–45.
- Kail, P. (2007). *Projection and Realism in Hume's Philosophy*. Oxford: Oxford University Press.



- Kail, P. (2016). Hume's 'Manifest Contradictions'. *Royal Institute of Philosophy Supplement*, 78, 147–160.
- Kampa, S. (2018). A New Statistical Solution to the Generality problem. *Episteme*, 15(2), 228–244.
- Loeb, L. E. (2002). *Stability and Justification in Hume's Treatise*. Retrieved from <http://dx.doi.org/10.1093/0195146581.001.0001>
- Loeb, L. E. (2006). Psychology, Epistemology, and Scepticism in Hume's Argument about Induction. *Synthese*, 152(3), 321–338.
- Loeb, L. E. (2008). Inductive Inference in Hume's Philosophy. In E. S. Radcliffe (Ed.), *A Companion to Hume* (pp. 106–125). Oxford: Wiley-Blackwell.
- Loeb, L. E. (2011). Epistemological Commitment in Hume's *Treatise*. *Oxford Studies in Early Modern Philosophy*, 6, 309–247.
- Lyons, J. (2019). Algorithms and Parameters: Solving the Generality problem for Reliabilism. *The Philosophical Review*, 128(4), 463–509.
- Marusic, J. S. (2010). Does Hume Hold a Dispositional Account of Belief? *Canadian Journal of Philosophy*, 40(2), 155–184.
- McAllister, B. (2018). Seemings as Sui Generis. *Synthese*, 195, 3079–3096.
- Meeker, K. (2013). *Hume's Radical Scepticism and the Fate of Naturalized Epistemology*. Basingstoke: Palgrave Macmillan.
- Millican, P. (1995). Hume's Argument Concerning Induction: Structure and Interpretation. In S. Tweyman (Ed.), *David Hume: Critical Assessments* (pp. 91–144). London and New York: Routledge.
- Millican, P. (2002). Hume's Sceptical Doubts concerning Induction. In P. Millican (Ed.), *Reading Hume on Human Understanding* (pp. 107–174). Oxford: Oxford University Press.
- Millican, P. (2012). Hume's 'Scepticism' About Induction. In A. O. B. Bailey, Dan (Ed.), *The Continuum Companion to Hume* (pp. 57–103). New York: Bloomsbury Academic.
- Millican, P. (2016). Hume's Chief Argument. In P. Russell (Ed.), *The Oxford Handbook of Hume* (pp. 82–108). New York: Oxford University Press.
- Mounce, H. (1999). *Hume's Naturalism*. London: Routledge.
- Noonan, H. (1999). *Hume on Knowledge*. London: Routledge.
- Owen, D. (1999). *Hume's Reason*. Retrieved from <http://dx.doi.org/10.1093/0199252602.001.0001>
- Pritchard, D. (2007). Wittgenstein's *On Certainty* and Contemporary Anti-Scepticism. In D. Moyal-Sharrock & W. H. Brenner (Eds.), *Readings of Wittgenstein's On Certainty* (pp. 189–224). Basingstoke: Palgrave Macmillan.

- Pryor, J. (2000). The Skeptic and the Dogmatist. *Noûs*, 34(4), 517–549.
- Pryor, J. (2004). What's Wrong With Moore's Argument? *Philosophical Issues*, 14(1), 349–378.
- Qu, H. (2014a). Hume's Positive Argument on Induction. *Noûs*, 48(4), 595–625. doi:10.1111/nous.12015
- Qu, H. (2014b). Hume's Practically Epistemic Conclusions? *Philosophical Studies*, 170(3), 501–524. doi:10.1007/s11098-013-0260-1
- Qu, H. (2016). Prescription, Description, and Hume's Experimental Method. *British Journal for the History of Philosophy*, 24(2), 1–23. doi:10.1080/09608788.2015.1134438
- Qu, H. (2017). Hume's Doxastic Involuntarism. *Mind*, 126, 53–92.
- Qu, H. (2018a). Hume's (*Ad Hoc?*) Appeal to the Calm Passions. *Archiv für Geschichte der Philosophie*, 100(4), 444–469.
- Qu, H. (2018b). Hume's Internalism in EHU 12. *Philosophy and Phenomenological Research*, 96(3), 517–539.
- Qu, H. (2020). *Hume's Epistemological Evolution*. New York: Oxford University Press.
- Sasser, N. (Forthcoming). Hume's Purely Practical Response to Philosophical Scepticism. *Hume Studies*.
- Schafer, K. (2014). Curious Virtues in Hume's Epistemology. *Philosophers' Imprint*, 14(1), 1–20.
- Schmitt, F. (1992). *Knowledge and Belief*. London: Routledge & Kegan Paul.
- Schmitt, F. (2014). *Hume's Epistemology in the Treatise*. Oxford: Oxford University Press.
- Shaw, D. (1989). Hume's Theory of Motivation. *Hume Studies*, 15(1), 163–183.
- Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell Publishing.
- Steup, M. (2004). Internalist Reliabilism. *Philosophical Issues*, 14(1), 403–425.
- Stroud, B. (1977). *Hume*. London: Routledge and Kegan Paul.
- Tang, W. H. (2016). Reliability Theories of Justified Credence. *Mind*, 125(497), 63–94.
- Teng, L. (2018). Is Phenomenal Force Sufficient for Immediate Justification? *Synthese*, 195, 637–656.
- Tolhurst, W. (1998). Seemings. *American Philosophical Quarterly*, 35(3), 293–302.
- Tolly, J. (2017). A Defence of Parrying Responses to the Generality Problem. *Philosophical Studies*, 174(8), 1935–1957.

- Tucker, C. (2010). Why Open-Minded People Should Endorse Dogmatism. *Philosophical Perspectives*, 24, 529–545.
- Van Cleve, J. (2003). Is Knowledge Easy—Or Impossible? Externalism as the Only Alternative to Skepticism. In S. Luper (Ed.), *The Sceptics: Contemporary Essays* (pp. 45–60). Aldershot: Ashgate.
- Vogel, J. (2000). Reliabilism Leveled. *Journal of Philosophy*, 91(1), 602–623.
- Werner, P. (2014). Seemings: Still Dispositions to Believe. *Synthese*, 191, 1761–1774.
- Winkler, K. (1999). Hume's Inductive Scepticism. In M. Atherton (Ed.), *The Empiricists: Locke, Berkeley, and Hume* (pp. 183–212). Oxford: Rowman and Littlefield Publishers.
- Wright, C. (2002). (Anti-)sceptics simple and subtle: Moore and McDowell. *Philosophy and Phenomenological Research*, 65(2), 330–348.
- Wright, C. (2004). On Epistemic Entitlement: Warrant for Nothing (and Foundations for Free?). *The Aristotelian Society, Supplementary Vol.* 78(1), 167–212.
- Wright, C. (2014). On Epistemic Entitlement (II): Welfare State Epistemology. In D. Dodd & E. Zardini (Eds.), *Scepticism and Perceptual Justification* (pp. 213–247). Oxford: Oxford University Press.



## THE B SIDE OF IMAGINATION. HUME ON IMPERFECT IDEAS

**Abstract:** *My aim is to look into the representational aspect of ideas, exploring not only to what Hume refers as adequate ideas, but also these cases where for a number of reasons an idea does not reach that standard. It has been suggested that the latter are fictions, but an in-depth examination of Hume texts reveals that there are several types of imperfections, such as incompleteness or imprecision that prevent an idea from being adequate. This leads to an analysis of the status of supposed or pretended ideas, and the possibility of there being terms with no ideas annexed to them.*

**Keywords:** *Adequate ideas, fictions, supposed ideas, empty terms, representation*

### Introduction

In his recent book, *The Imagination in Hume's Philosophy* (2018), Timothy Costelloe offers an organized and systematic view about the workings of imagination in all of Hume's writings. Throughout the book, a distinction emerges between fictions and genuine ideas (2018, 30) which guides many of Costelloe's arguments.

The distinction develops as a result of examining two different powers of imagination: the mimetic and the productive (2018, 6). The mimetic power is related to the capacity of imagination to copy impressions in a similar way as memory does (T 1.1.3.1; SBN 8–9, EHU 2.2; SBN 17–18),<sup>1</sup> which provides reason and understanding the ideas they work with (2018, 6–7). The productive power, Costelloe argues, belongs to imagination understood in a narrow sense, that is, as excluding demonstrative and probable reasoning (T 1.3.9.19 n22, SBN 117–118). It has two sides: one combinatory and another creative. The former exhibits the power of imagination to combine, separate, rearrange, and change its stock of ideas in all possible ways to create new ideas (T 1.1.3.4, SBN 10; EHU 2.5, SBN 19; 5.10, SBN 47–48; 5.12, SBN 48–50). The latter refers to the capacity to create original ideas independently of experience and according to imagination's inner logic (2018, 21).<sup>2</sup> Thus, the

---

1 All works by Hume will be cited using abbreviations referenced in the Bibliography.

2 Costelloe states that the creative aspect of imagination is revealed when Hume argues that anyone can produce the idea of the missing shade of blue even though "it had never been convey'd to him by his senses" (T 1.1.1.10, SBN 5–6). Costelloe observes that the

productive aspect of imagination reflects the possibility to conceive objects that cannot be traced directly to an impression (2018, 25).

According to Costelloe, ideas that do not have a direct correlation with impressions spring up from certain operations of imagination, which slides easily from experience to what Hume sometimes refers to as „suppositions,“ i. e. something that is taken as if it were the case even though it has not been perceived or lacks justification (2018, 24). As a consequence of this easy transition, the imagination produces errors (2018, 25). Costelloe regards fictions as a type of error formed through the productive aspect of imagination (2018, 29).

Costelloe's distinction between genuine ideas and fictions is suggestive because it focuses not only on full-blooded ideas, but also on what does not quite fit in that place. However, when reviewing Hume's texts, it is not clear that the distinction between genuine ideas and those that are not genuine is exhaustive. Instead, we find a series of names for ideas that do not reach a desired standard: non-adequate, imperfect, supposed or pretended ideas. Despite what Costelloe argues, this does not necessarily entail all these ideas are errors, nor that they are detached from matters of fact either. It shows that for various reasons, they do not reach the status of genuine ideas. And the reasons offered to show that an idea is not genuine are different in each case, what lead us to suspect they cannot be considered fictions alike. Hume, as usual, does not present this sort of taxonomy of ideas systematically; hence difficulties arise when trying to reconstruct it in an organised way.

The different types of “imperfections” which prevent an idea from being genuine merit a deeper investigation. I first explain what makes an idea genuine, and then turn to those cases where these requirements cannot be met. This reveals that the differences between the various types of imperfect ideas are related to several representational problems between ideas and their objects. Finally, I trace some of the consequences of these problems at a semantic level, by means of examining the connection between imperfect ideas and ambiguous terms.

## Adequate ideas

First of all, I want to make clear that the expression „genuine idea“ is Costelloe's and not Hume's. When Hume refers to a full-blooded idea, he calls it „adequate idea“ instead. Therefore, I will turn to Humean terminology and proceed to study the nature of these perceptions.

The notion of adequate idea or adequate knowledge was common in Early Modern philosophy. Roughly speaking –although there are differences among authors such as Descartes, Leibniz, Locke or Spinoza– it refers to adequate representations of objects, in the sense that any attribute of an adequate idea is also an attribute of the object it represents (Lightner 1997, 117). Hume does

---

creative aspect of imagination has not been noticed by commentators, with the exception of Donald Ainslie (2015, 64).

not explicitly define what he means by an adequate idea, and there are some nuances expressed by the interpreters in this regard. Donald Ainslie (2010, 41; 2015, 69) relates Hume's notion of adequate idea to that of Descartes and Locke, understanding it as an idea that „fully capture [its] object[s].“ In a similar vein, Peter Kail (2003, 52–53), argues that, in line with Locke, Hume considers that adequate ideas offer all the relevant information of an object, in the sense of reflecting its internal structure or complete essence. Tyckerium Lightner (1997, 121), on the other hand, states that Hume uses “adequate,” to mean “clear and distinct.” This, in fact, is not incompatible with the definition proposed by Ainslie and Kail, since an idea that fully captures its object is, after all, a clear and distinct idea. As I will show below, the adequate character of an idea resides not only in its completeness (i. e. that it can fully represent its object) but also in its precision (i. e. that it is able to represent its object's attributes in a distinct and precise way).

When defining adequate ideas, scholars frequently rely on the passage where Hume states: “Wherever ideas are adequate representations of objects, the relations, contradictions and agreements of the ideas are all applicable to the objects” (T 1.2.2.1, SBN 29). Setting aside the various interpretations this passage has been subject to, I wish to remark how the passage shows that, like other modern philosophers, Hume relates adequacy to the representational ability of an idea. Of course, determining what Hume intends as an object would merit a separate work. I restrict myself to pointing out that if by an object we understand the impression(s) that the idea copies, then an idea could be an adequate representation of its object. But as Kail (2003, 56n23) observes, if by an object we mean an *in re* object, we cannot have an adequate impression of it, nor therefore an adequate idea, since we cannot conceive of it as something specifically different from a perception (T 1.2.6.8, SBN 67–8).

Hume touches on adequate ideas in the *Dialogues concerning Natural Religion* when considering the possibility of knowing divine nature and attributes. This text has been overlooked by scholars who dealt with the subject, but it offers extra evidence to clarify the features of completeness and precision I mentioned above. Cleanthes aims at attaining an adequate idea of the Divine mind from its supposed similarity to the human mind. By means of the argument from design, he transfers the similarity between the products of human contrivance and the order of the universe, to a supposed similarity between their causes. Nevertheless, from Demea's mystical perspective, human nature is incommensurable with Divine nature. This means that there is no way of knowing whether any feature of the former can be similar to the latter so as to achieve, by means of analogy, an adequate idea of Divinity:

Your instance, Cleanthes, said he [Demea], drawn from books and language, being familiar, has, I confess, so much more force on that account; but is there not some danger too in this very circumstance; and may it not render us presumptuous, by making us imagine we

comprehend the Deity, and have some adequate idea of his nature and attributes? When I read a volume, I enter into the mind and intention of the author: I become him, in a manner, for the instant; and have an immediate feeling and conception of those ideas, which revolved in his imagination, while employed in that composition. But so near an approach we never surely can make to the Deity. His ways are not our ways. His attributes are perfect, but incomprehensible (DNR 3.11, KS 155–6).

Cleanthes' argument relies on one of the relevant conditions for an idea to represent an object, which is resemblance (Garrett 2006, 308). In this case we cannot have an impression of the Divine attributes for obvious reasons, but Cleanthes believes that this difficulty can be overcome by appealing to the similarity between their effects and those belonging to a mind whose impression we do have: the human. This means that if we can have an adequate idea of the human mind from the works of its contrivance, we can have an adequate idea of the divine mind from His creation. Demea, in turn, argues against the legitimate character of this similarity. Don Garrett (2006, 309) points out that, in addition to resemblance, there is a second condition that amounts to representation. This condition is causal derivation, since Hume affirms that "ideas always represent the objects or impressions from which they are derived" (T 1.2.3.11, SBN 37). Demea's mystical perspective would allow God being the immediate cause of the idea we have of Him (since, again, we can have no impression of that idea), but this possibility does not seem to grant that our idea represents His nature adequately. Unlike causal derivation, resemblance seems to be in this case a crucial condition to entertain an adequate idea of a certain object. As long as there is nothing in the divine attributes that resembles ours, it is not possible to attain an adequate idea of them, because there are no means of establishing a comparison between both. Cleanthes, on the other hand, claims that even though we can establish a limited similarity only, this amounts to an adequate idea of God:

It seems strange to me, said Cleanthes, that you, Demea, who are so sincere in the cause of religion, should still maintain the mysterious, incomprehensible nature of the Deity, and should insist so strenuously, that he has no manner of likeness or resemblance to human creatures. The Deity, I can readily allow, possesses many powers and attributes, of which we can have no comprehension: But if our ideas, so far as they go, be not just, and adequate, and correspondent to his real nature, I know not what there is in this subject worth insisting on. Is the name, without any meaning, of such mighty importance? (DNR 4.1, KS 158)

Cleanthes admits that the similarity between God and His creatures is not complete. But this fact does not seem to preclude us from having an adequate idea of Him, that is, an idea that corresponds to His „real nature.“ Even



though there are many divine attributes that exceed our cognitive capacity, the resemblance between certain Divine and human attributes has been legitimately demonstrated by the argument from design. This grants that our idea of God is adequate. From Cleanthes's perspective, then, an idea that is admittedly incomplete, since it does not "fully capture [its] object," can be adequate, after all; whereas from Demea's point of view, if an idea is incomplete, it cannot be adequate.

So far, the *Dialogues* leave us with two possible ways of understanding an adequate idea, reflected in Demea's and Cleanthes' positions. There is no way of being certain about which of the characters represent Hume's position, or even if it is represented by any of them at all. However, they present us with a spectrum of possibilities that can be confronted with other texts to see if Hume assumes some of their features as his own explicit view. Let us turn to the *Treatise* in order to do so:

First then I observe, that when we mention any great number, such as a thousand, the mind has generally no adequate idea of it, but only a power of producing such an idea, by its adequate idea of the decimals, under which the number is comprehended. This imperfection, however, in our ideas, is never felt in our reasonings (T 1.1.7.12, SBN 22–23).

In this passage, Hume seems to understand adequate ideas in terms of representations that fully capture their objects, in line with Demea's perspective. But unlike the idea of God, in this case the object represented falls within the scope of experience. This grants we can have an impression of it. Thus, the difficulty in achieving an adequate representation does not lie in the incommensurability between the object and the idea that stands for it. The idea of a thousand is not adequate because it is difficult to verify whether the units we envision actually reach that amount. However, Hume claims that we can generate an approximate and imperfect representation of a great number with the help of an adequate idea. The adequate idea of a decimal, when replicated a number of times, allows us to understand the notion of a thousand, since the units that make up a certain number, be it ten, be it a thousand, are identical. We can notice, then, that inadequate ideas are not necessarily errors. As Ainslie argues (2015, 69), they may be sufficient for some purposes; such as to represent a large amount. If needed, we are endowed with the cognitive skills to be more precise and generate an idea that represents a thousand units.

Later on, Hume offers another example concerning the difficulty of having adequate ideas of extremely small animals such as a mite. It is important to notice that the previous example belongs to mathematics, where ideas are always „clear and determinate“ and have univocal or at least stable definitions. This sort of definitions can help us verify, eventually, the completeness of the representation (EHU 7.1, SBN 60). The mite example belongs to matters of fact, where we are deprived of this possibility:

This however is certain, that we can form ideas, which shall be no greater than the smallest atom of the animal spirits of an insect a thousand times less than a mite: And we ought rather to conclude, that the difficulty lies in enlarging our conceptions so much as to form a just notion<sup>3</sup> of a mite, or even of an insect a thousand times less than a mite. For in order to form a just notion of these animals, we must have a distinct idea representing every part of them; which (...) is extremely difficult, by reason of the vast number and multiplicity of these parts (T 1.2.1.5, SBN 28).

In this passage Hume points out that in order to have an adequate idea of minute creatures or objects, we must be able to conceive distinct ideas of each and every one of its parts, which is „extremely difficult.“ The idea of a mite is an extremely complex idea, because it is made up of a large amount of simple ideas. Even though imagination is able to form these kinds of ideas, the difficult thing is to verify whether we include not only all the parts that make it up, as with the idea of great numbers, but also, that we have a distinct idea of each of them. Unlike the concept of a thousand, we cannot formulate a definition that exhaustively describes the parts of a mite, let alone suppose that all of them are identical, which would allow us to suppose that by knowing some of its parts, we can have an idea of the whole, as in the case of the units that make up a great number. Nothing prevents us from thinking that a scientist may discover unknown parts of this animal using a powerful microscope (see T 1.2.1.4, SBN 27–8). In this case, thus, the idea is not adequate not only because it is incomplete but also because it is imprecise.

So far, textual evidence suggests that adequacy consist in two aspects: precision and completeness. If we guarantee the ideas we have are distinct then, although we cannot know for sure whether they fully reflect their object, we may reach a limited adequacy of the kind Cleanthes proposes. If we also guarantee that the idea is complete, we can have a fully adequate idea, which matches Demea’s requisites. But as the examples from the *Treatise* suggest, the two aspects of adequacy can only be achieved by –or at least, be verified in– objects belonging to the realm of relations of ideas, simple ideas of sensation or ideas of relatively low complexity.

Up to this point we can see that one of the main problems concerning adequate ideas consists in what Ainslie names the “epistemological challenge” (2010, 41). Ainslie argues that, unlike Descartes and Locke, Hume claims that adequacy of ideas does not refer to *in re* objects but to perceptual objects (2010, 61). However, even if we admit that the representational character of an idea refers to a perceptual object, this fact does not guarantee *per se* the possibility of fully determining their adequacy. The methodological use of

---

3 Although Hume is using here the term “just”, at the beginning of the paragraph he talks about forming “adequate ideas”, so at least in this case he is using “just” as a synonym for “adequate.”

the copy principle<sup>4</sup> can only solve the problem in part, because it allows us to deal with issues related to precision, but not to completeness. Hume asserts that impressions are always clear and distinct, and therefore incontrovertible, while ideas are usually obscure because they are weaker or fainter than impressions. As is well known, the copy principle allows us to clarify ideas by showing their connection with the impressions that caused them:

By what invention can we throw light upon these ideas, and render them altogether precise and determinate to our intellectual view? Produce the impressions or original sentiments, from which the ideas are copied. These impressions are all strong and sensible. They admit not of ambiguity. They are not only placed in a full light themselves, but may throw light on their correspondent ideas, which lie in obscurity. And by this means, we may, perhaps, attain a new microscope or species of optics, by which (...) the most minute, and most simple ideas may be so enlarged as to fall readily under our apprehension, and be equally known with the grossest and most sensible ideas, that can be the object of our enquiry (EHU 7.4, SBN 62).

The copy principle begins with ideas and proceeds backwards, towards their origin, by means of analysis or decomposition. This proceeding can concede greater precision to ideas of minute animals such as a mite, „enlarging“ our complex ideas to achieve a precise and distinct vision of the simple ideas that compose it, relating these simple ideas to the impressions that caused them. But the copy principle cannot reveal if the complex idea of the mite fully captures its object, i.e., it does not show if there are any impressions left; it just helps us realize if the ideas we envision are clear and precise. This problem, as indicated by Kail (2003, 49–50) and Ainslie (2010, 43) is reminiscent of what Locke says concerning the inadequacy of our ideas of substances, i. e. that it is impossible for them to include representations of all their qualities (E 2.31.11).<sup>5</sup> Although, unlike Locke, Hume denies that our ideas can represent anything different from perceptions, the possibility that our complex ideas do not „fully capture their objects“ persists. He himself suggests it when he says: “I have seen Paris; but shall I affirm I can form such an idea of that city, as will perfectly represent all its streets and houses in their real and just proportions?” (T 1.1.1.4, SBN 3). Apparently, then, there is no way to fulfil the completeness requirement, at least in the case of complex ideas concerning matters of fact. This brings Hume’s position closer to Cleanthes’s than to Demea’s: his references to adequate ideas should be understood in a limited sense.

---

4 I refer to the methodological use of the copy principle following Garrett’s suggestion (1997, 73), in the sense that it can be used as “the support for a methodological directive concerning problematic ideas.”

5 Locke’s work will be cited using abbreviations referenced in the Bibliography.

## Fictions and supposed ideas

We can move now to those ideas Hume regards as “imperfect” (T 1.1.7.12, SBN 22–3) in order to study why they fall short of being adequate. In the introduction I pointed out that fictions are one type of imperfect ideas, but not the only one. Let us see, in the first place, why at least certain kinds of fictions can be considered imperfect ideas, and then examine the case of supposed or pretended ideas.

As some authors point out (Traiger 1987, 395; Ainslie 2015, 83), Hume does not have a completely negative view of fictions, but rather of certain fictions that philosophers construct and apply, such as occult qualities or sympathies and antipathies. There are other fictions, which have been called „natural“ (Traiger 1987, 395; Costelloe, 2018, 33–4) because they arise from the regular motions of imagination, such as the easy transition principle. Among them we can mention personal identity, substance, the distinct and continuous existence of objects, duration of unchanging objects, and perfect standards, to name a few. These natural fictions fulfil certain epistemic and practical functions relevant to the development of everyday life. Philosophical fictions, on the contrary, fall within what Costelloe (2018, 31–2) calls “artificial fictions,” which are the product of an individual’s imagination.<sup>6</sup> They stem from a voluntary act of the imagination that can combine and create new ideas that do not directly correspond to matters of fact and existence.

Even though different types of fictions can be distinguished in Hume’s texts, they all share a distinct feature: they are ideas applied to something from which they could not have been derived (Traiger 1987, 385–386; Ainslie 2015, 66; Costelloe 2018, 30). Thus, Hume claims: “Ideas always represent the objects or impressions, from which they are deriv’d, and can never without a fiction represent or be apply’d to any other” (T 1.2.3.11, SBN 37). A few lines before he states that, with respect to the idea of duration applied to an unchanging object, „it can never in any property or exactness be applied to it” (T 1.2.3.11, SBN 37). As Saul Traiger notices (1987, 386), in this passage Hume relates the imperfect character of a fiction to its being improper, because it arises in one context and is improperly applied to another. The idea of duration is never derived from an unchanging object, but from the comparison between that object and the incessant flow of our perceptions. The fiction consists of applying that idea, which derives from constant movement, to a static object. Ainslie (2015, 75) calls this the “idea-substitution mechanism,” and describes it as an operation of the imagination where an idea is inserted in the place of another without our realizing it. As I mentioned above, Hume does not hold a negative view of this mechanism, because he regards it as natural. What Hume condemns is that certain philosophers omit this substitution mechanism carried out by

6 Costelloe (2018, 29–36) develops a detailed analysis of the different types of fictions.

the imagination and assume that the ideas in question are indeed derived from such objects (Traiger 1987, 386) or else they intend to apply them to non-perceptual objects (Ainslie 2015, 83, 106).

The reason why Hume considers (at least some) fictions as imperfect ideas is different from the ones mentioned in the previous section. The decisive feature in this case is not imprecision or incompleteness, but impropriety, that is, the fact that the idea supposedly represents an object that it cannot represent. This means that a fiction does not meet the requirement of causal dependence: it could not have been derived from the object we improperly intend to apply it. If we recall the relevant conditions that amount to the representational character of an idea, we can relate completeness and precision to resemblance, and property to causal dependence. When any of these conditions fails, the result is an imperfect idea.

In addition to the fictions already mentioned, Garrett (2006, 313 n17) recognizes another type of fictions he considers more philosophically objectionable: “we take ourselves to have ideas meeting certain representational conditions –in the case of ‘substance’ for example, the condition of representing something that unites complex and interrupted items into something simple and identical– without having any idea at all (T 1.1.6.1; T 1.4.3).” Nevertheless, I think that regarding this case as a type of fiction is troublesome; because it entails assuming we have an idea when in fact we have no idea at all, and Hume never calls into question that fictions are ideas (T 1.3.7.7, SBN 628–29; EHU 5.10–11, SBN 47–8). Let us briefly review the two passages Garrett mentions as textual evidence of this type of fiction. This is the first one:

I would fain ask those philosophers, who found so much of their reasonings on the distinction of substance and accident, and imagine we have clear ideas of each, whether the *idea* of substance be deriv'd from the impressions of sensation or of reflection? (...) We have therefore no idea of substance, distinct from that of a collection of particular qualities, nor have we any other meaning when we either talk or reason concerning it (T 1.1.6.1, SBN 15–16).

It can be noticed that Hume does not question there is an idea of substance. Rather, he intends to show that it is a fiction in the sense defined above. The notion of substance is nothing but a collection of simple ideas united by the imagination, but philosophers pretend to refer it “to an unknown something, in which they [the qualities] are supposed to inhere; or granting this fiction should not take place, are at least supposed to be closely and inseparably connected by the relations of contiguity and causation” (1.1.6.2, SBN 16). In other words, this is an instance similar to the other cases of fictions we have seen. Then, Garrett sends us to 1.4.3, where Hume continues talking about substance. We find this passage:

But these philosophers carry their fictions still farther in their sentiments concerning occult qualities, and both suppose a substance supporting, which they do not understand, and an accident supported, of which they have as imperfect an idea (T 1.4.3.8, SBN 222).

Here, Hume regards both substances and occult qualities as fictions, which in turn he considers as imperfect ideas.<sup>7</sup> This shows that a fiction never loses its status of idea. If we assume we have an idea when in fact we cannot be sure we have it, then we are not entertaining a fiction, but something else. Garrett is correct in acknowledging the existence of this type of philosophically objectionable representational items, but I do not think he is right in classifying them as fictions. These items are what Hume calls „supposed“ or „pretended“ ideas, and what I will examine now is whether they are indeed some other type of imperfect ideas or they are something else, in case we do not have „any idea at all.“

Hume mentions these items for the first time in the *Abstract* and then again in *An Enquiry concerning Human Understanding*. Both passages have a similar structure. In the former, he refers to them as „pretended ideas.“<sup>8</sup> The notion appears when Hume is introducing the methodological use of the copy principle:

Our author thinks, that no discovery cou'd have been made more happily for deciding all controversies concerning ideas than this, that impressions always take the precedency of them, and that every idea, with which the imagination is furnish'd, first makes its appearance in a correspondent impression. These latter perceptions are all so clear and evident, that they admit of no controversy; tho' many of our ideas are so obscure, that 'tis almost impossible even for the mind, which forms them, to tell exactly their nature and composition. Accordingly, wherever any idea is ambiguous, he has always recourse to the impression, which must render it clear and precise. And when he suspects that any philosophical term has no idea annexed to it (as is too common) he always asks from what impression that pretended

---

7 Admittedly, the idea of substance is a particularly difficult case. At T 1.4.5.6, SBN 234, Hume states: “We have no perfect idea of any thing but of a perception. A substance is entirely different from a perception. We have, therefore, no idea of a substance.” Although it might seem that he actually affirms there is no idea of substance –which on the other hand would contradict T 1.1.6.1– the context shows that Hume is referring to the philosophical notions of substance as an inhesion substrate, or as something that may exist by itself. David Landy (2018, 107 n5) argues that Hume is not criticizing the bare notion of substance, but a particular idea of substance „as incorrectly conceived by his predecessors (e. g. Cartesians).“ We could then suppose that what Hume means is that there is no perfect idea of substance, since its fictitious character entails it is an imperfect idea. This interpretation makes T 1.4.5.6 consistent with what he claims at T 1.4.3.8.

8 Given the similarities between the two passages, I assume the adjectives „supposed“ and „pretended“ can be considered equivalent.

idea is derived? And if no impression can be produced, he concludes that the term is altogether insignificant. 'Tis after this manner he examines our idea of substance and essence; and it were to be wished, that this rigorous method were more practised in all philosophical debates (Abs. 7, SBN 648–9).

In this paragraph we can distinguish two levels: the mental and the linguistic (Ott 2009, 198). At the mental level, Hume deals with the representational relation between ideas and impressions, which can be clarified using the copy principle. At the linguistic level, he deals with the relation between words and ideas, which he describes in terms of annexation. If we assume that the copy principle can trace the origin of an idea, then sooner or later we must find the impression(s) that caused it. This is what Hume states at the beginning of the paragraph: every [simple] idea has as its antecedent a [simple] impression. So, this seems to rule out the possibility that an idea does not originate, directly or indirectly, in an impression.<sup>9</sup> However, the text does not seem to indicate that the relationship between words and ideas has the same universal character that unites ideas with impressions.<sup>10</sup> Therefore, if the existence of a word does not necessarily indicate the existence of an idea, it would be possible that terms without attached ideas exist. We might consider these as „empty“ terms or, as Hume calls them, „insignificant.“ A word means an idea, and if there is no such idea, then the word is meaningless. This situation can be verified turning to the copy principle, which cannot directly clarify the relationship between the word and the pretended idea, but rather between the pretended idea and the impression(s) from which it could have been derived. If we cannot find the impression(s) that might have caused the pretended idea, then it seems that the term in question does not have any correlation at the mental level. It is nonsense (Ott 2009, 208–9).

Now, it is important to note that Hume states the proposition where he deals with words-ideas relationship in a conditional mode, since he qualifies it as a suspicion. This means he opens both the possibility of not finding or producing the impression(s) from which that pretended idea would have been derived, as well as the possibility of finding it, which would modify its pretended status for that of a plain idea. This suggests that the category of pretended idea does not necessarily mark a lack of ideas but rather directs our attention to a point that requires clarification. As a result, the examination of pretended ideas can yield an idea (be it adequate or imperfect, depending on the relation it has with the impression(s) found), or no idea at all, in the event

---

9 Several authors put this strict interpretation of the copy principle into question and remark the creative capacity of the imagination to produce ideas detached from impressions. As stated in footnote 2, among them are Ainslie and Costelloe. Also, the New Humeans propose a broad interpretation of the theory of ideas based on the category of relative ideas (Richman 2000, 3–4). However, in the passages where Hume mentions supposed and pretended ideas, he seems to understand the copy principle in a strict sense.

10 I will return to the semantic issue in the last section.

that we cannot produce any impression that directly or indirectly caused it. The allusion to substance at the end of the paragraph confirms the former possibility: in this case, the analysis would not lead us to an empty term, as Garrett's interpretation suggests, but to a fictitious idea.

However, there is something awkward in this interpretation of the passage from the *Abstract*. Hume seems to be applying the copy principle to something whose ideal status is in doubt and may consist of just an empty linguistic term. Would this use of the copy principle be admissible? Should not we be sure to apply the copy principle to an idea, be it obscure or faint, but an idea nonetheless? Let us turn to the second passage, which belongs to the first *Enquiry*, to see if it helps us clarify this matter. Here, Hume refers to these items as "supposed ideas":

All ideas, especially abstract ones, are naturally faint and obscure: The mind has but a slender hold of them: They are apt to be confounded with other resembling ideas; and when we have often employed any term, though without a distinct meaning, we are apt to imagine it has a determinate idea, annexed to it. On the contrary, all impressions, that is, all sensations, either outward or inward, are strong and vivid: The limits between them are more exactly determined: Nor is it easy to fall into any error or mistake with regard to them. When we entertain, therefore, any suspicion, that a philosophical term is employed without any meaning or idea (as is but too frequent), we need but enquire, from what impression is that supposed idea derived? And if it be impossible to assign any, this will serve to confirm our suspicion. By bringing ideas into so clear a light, we may reasonably hope to remove all dispute, which may arise, concerning their nature and reality (EHU 2.9, SBN 21–2)

This passage preserves in outline the structure of the analogous one from the *Abstract*. There are certain key points that appear in both. First, the "supposed" or "pretended" character is related to the faint and obscure nature of ideas, which, in turn, gives rise to an ambiguity that causes a series of problems. The ambiguity inherent in many ideas –expressed in the fact that some ideas can be confounded with similar ones– is transferred to the terms we attach to them. Second, Hume does not claim that the impressions that could have originated this supposed idea do not exist, but rather that we cannot "produce them" or "assign any," that is, we cannot find them. This implies that, while the situation persists, we cannot clarify the supposed idea and therefore we cannot disambiguate the term to which the idea is annexed. Thus, Hume might be suggesting that the term is not strictly empty but rather attached to an idea that remains unclear for the moment. In the passage from the first *Enquiry* this last option seems to be more forceful than that of the empty term. Certain qualifications Hume employs seem to direct us this way. In the first part, he says we use a term „without any distinct meaning,“ and we imagine it has annexed a „determinate idea“ to it. The use of "distinct" and



“determinate” indicates that the term lacks a precise meaning, rather than has no meaning at all, and that the idea attached to it is faint and obscure, but it is an idea after all. In the second part of the paragraph he takes up the same proposition although omitting the qualifiers „distinct“ and „determinate.“ However, given their appearance in the first half of the paragraph it is highly plausible that Hume is referring to these kinds of terms and ideas. I suggest the sentence should be read along these lines: “When we entertain, therefore, any suspicion, that a philosophical term is employed without any [distinct] meaning or [determinate] idea (as is but too frequent), we need but inquire, from what impression is that supposed idea derived?”<sup>11</sup>

The text from the *Enquiry* seems to lead us towards a moderate interpretation of supposed ideas, since it suggests they are ideas and not empty terms. We can regard them as an extreme case of imperfect ideas, because they remain in a state of confusion or indistinctness, although this situation is not necessarily definitive. One of the advantages of this interpretation is that it preserves the regular use of the copy principle: if we consider that supposed ideas are ideas after all, the copy principle would apply to the relationship between ideas and impressions without forcing it into the linguistic level, a move which, on the other hand, would be difficult to justify.

## Fictions vs. supposed ideas

If we give preference to the moderate interpretation of supposed ideas and admit they are ideas after all, why cannot they be considered fictions, as Garrett suggests? Before answering this question it is worth noticing, in the first place, that the criticism I made of Garrett’s classification still stands because he poses that something consisting in „no idea at all“ can be a fiction. I provided textual evidence to show that this option is not plausible because fictions are (imperfect) ideas. In the second place, I have also shown that an idea can be imperfect for a number of reasons. Therefore, the fact that supposed ideas are imperfect does not necessarily imply they are fictions. What is the difference between a supposed idea and a fiction, then?

To begin, remember that the imperfect character of a fiction resides not in its incompleteness or imprecision, but in its impropriety: we consider that it represents something it could not have been copied from. Throughout the *Treatise* we find many examples of this kind, where Hume shows the real origin of these fictitious ideas and explains why it is improper to apply them to another context. I already mentioned the idea of duration without change and of substance, but other instances can be added: the perfect standard of equality (T 1.2.4.24, SBN 47–9), the distinct and continuous existence of

---

11 We would be facing a case similar to that of the idea of substance mentioned in footnote 7. When Hume states the proposition for the second time, he omits the qualifier “perfect” when referring to that idea, and says: “we have, therefore, no idea of a substance.”

objects (T 1.4.2), or the Self (T 1.4.6), to mention a few more. A supposed idea, on the other hand, is not an improper idea, but a radically imprecise idea, the obscurity of which cannot be dissipated so far. Let us review a couple of passages where Hume seems to suggest the difference between fictions and supposed ideas along these lines. In the first one, he is discussing the fiction of duration without change:

If it be a sufficient proof, that we have the idea of a vacuum, because we dispute and reason concerning it; we must for the same reason have the idea of time without any changeable existence; since there is no subject of dispute more frequent and common. But that we really have no such idea, is certain. For whence shou'd it be deriv'd? Does it arise from an impression of sensation or of reflection? Point it out distinctly to us, that we may know its nature and qualities. But if you cannot point out *any such impression*, you may be certain you are mistaken, when you imagine you have *any such idea*.

But tho' it be impossible to shew the impression, from which the idea of time without a changeable existence is deriv'd; yet we can easily point out those appearances, which make us fancy we have that idea. (T 1.2.5.28–29, SBN 64–5).

The first paragraph proceeds in the usual way, pointing out to the relationship between linguistic and mental levels. Hume argues that the fact of having the terms „vacuum“ and „time without change“ does not indicate we have adequate ideas of them. As in EHU 2.9, he omits the qualification “adequate” and simply says that „we have no such idea.“ So far, the procedure is similar to the one we saw at Abs. 7 and EHU 2.9. However, after suggesting the use of the copy principle to solve the issue, he says that „we can easily point out those appearances which make us fancy we have that idea.“ This means that, unlike supposed ideas, we can explain why the idea of duration without change is imperfect and by what means it was created. In the case of supposed ideas, there would be no explanation available.

Let us examine another example. In this case, Hume is dealing with the idea of necessary connection:

when we speak of a necessary connexion betwixt objects, and suppose, that this connexion depends upon an efficacy or energy, with which any of these objects are endow'd; in all these expressions, *so apply'd*, we have really no distinct meaning, and make use only of common words, without any clear and determinate ideas. But as 'tis more probable, that these expressions do here lose their true meaning by being *wrong apply'd*, than that they never have any meaning; 'twill be proper to bestow another consideration on this subject, to see if possibly we can discover the nature and origin of those ideas, we annex to them (T 1.3.14.14, SBN 162).

Again, we find the same structure. Hume recognizes there is a term, „necessary connection“, which represents an idea he considers imperfect (because it is not „clear and determinate“). Given this situation, we are presented with two options: either it is a term that has no distinct meaning (and we again notice the ellipsis of the qualifier „distinct“ when he repeats the sentence, since it says „never have any meaning“), an option that would lead us to supposed ideas, or it lost its true meaning due to having been misapplied, which would mean the idea of necessary connection is a fiction. Hume believes we are faced with this last option and proceeds to explaining the mechanism that produced the fiction below. In presenting the disjunction, Hume says it is more likely that when we come across these problematic cases, we are dealing with fictions and not with supposed ideas. Is there any text where Hume not only suggests but also asserts that we are faced with ideas in a state of radical ambiguity? In the *Dialogues*, Philo presents an instance of this sort when discussing about degrees of qualities and circumstances:

But there is a species of controversy, which, from the very nature of language and of human ideas, is involved in perpetual ambiguity, and can never, by any precaution or any definitions, be able to reach a reasonable certainty or precision. These are the controversies concerning the degrees of any quality or circumstance (...) The disputants may here agree in their sense, and differ in the terms, or vice versa; yet never be able to define their terms, so as to enter into each other's meaning: Because the degrees of these qualities are not, like quantity or number, susceptible of any exact mensuration, which may be the standard in the controversy. That the dispute concerning Theism is of this nature, and consequently is merely verbal, or perhaps, if possible, still more incurably ambiguous, will appear upon the slightest enquiry (DNR 12.7, KS 217–8).

In this case, the radical ambiguity of ideas stem from the impressions they copy, because the latter cannot be expressed in precise units of measure. Philo says we can dispute endlessly „whether Hannibal be a great, or a very great, or a superlatively great man, what degree of beauty Cleopatra possessed, what epithet of praise Livy or Thucydides is intitled to, without bringing the controversy to any determination“ (DNR 12.7, KS 217). In the *Treatise* Hume argues that we can intuitively recognize the degrees of a secondary quality such as weight, heat, colour or flavour when the difference between the compared objects is large (T 1.1.5.7, SBN 15; 1.3.1.3, SBN 70) and even considers, as mentioned before, the fiction of a perfect standard for quantifiable items. But in the *Dialogues*, he is not referring to these kinds of qualities, but rather to others of a moral and aesthetic nature; and mainly to circumstances where we are comparing something known –the human mind– to something unknown –the divine mind (DNR 7.12, KS 218). This feature renders ideas that

represent these kinds of qualities and circumstances imprecise and prevents reaching an agreement on their definition. If achieved, this agreement would guarantee the strict and uniform use of the terms we attach to those ideas.<sup>12</sup>

## Words and ideas

Finally, I am going to deal briefly with the relation of linguistic and mental levels, which, as has become apparent, brings about consequences in the distinction between adequate and imperfect ideas. The difficulty in addressing this question is that Hume never systematically developed a theory of language (Powell 2013, 112; Ainslie 2010, 50). However, in general terms, it can be said that for Hume, every word with meaning must be related to an idea (Ott 2006, 235). We have already seen that Hume states ideas are „annexed“ to words. But in the light of what has been discussed, it is necessary to determine, firstly, whether meaning comes from an adequate idea only, or imperfect ideas can confer at least a minimum of intelligibility on the terms they are annexed to. Secondly, I will deal with the nature of the relation between words and ideas in order to elucidate whether it is indeed possible that there are “empty” terms strictly speaking, that is, detached from any idea, or, as I proposed earlier, these are terms connected with supposed ideas.

Regarding the first question, it can be contended that adequate ideas confer precision and uniformity to the terms connected with them, while obscure and imprecise ideas transfer their ambiguity to the terms they are attached to (EHU 7.2, SBN 60–1). Usually, Hume points out that ambiguity is greater or incurable when we deal with subjects that exceed our cognitive capacity (EHU 8.1, SBN 80–1), hence metaphysics is the usual domain where this kind of problems appears. An obvious case is that of the divine nature and attributes we saw earlier. But I also noted above that issues regarding degrees of certain empirical qualities, such as the greatness of Hannibal or the beauty of Cleopatra, seem hard to disambiguate. Furthermore, in the first *Enquiry* Hume indicates that this also takes place in the moral sciences because their objects are difficult to perceive: “the finer sentiments of the mind, the operations of the understanding, the various agitations of the passions, though really in themselves distinct, easily escape us, when surveyed by reflection” (EHU 7.1, SBN 60).

Hume suggests this kind of problem can be solved by „fixing“ the meaning of ambiguous terms. This amounts to clarifying obscure ideas

---

12 Moral and aesthetic standards are particularly difficult to establish because they are grounded in feeling. Angela Coventry (2006, 120–133) argues it is possible to construct an ideal standard in these realms if we have a delicate imagination and go through a long process of observation, review, discussion and correction. This example shows that the state of radical ambiguity of supposed ideas is not definitive, since it is possible to find a way of clarifying them.

attached to them via the copy principle. However, as I already noted, the use of the copy principle is limited. Consequently, Hume admits that we must not take for granted its application would lead us invariably to precise and uniform terms. That is why he is very cautious when describing his intentions: “We shall endeavour (...) to fix if possible the meaning of these terms, and thereby remove some part of that obscurity, which is so much complained of in this species of philosophy” (EHU 7.3, SBN 61–2). Why be so cautious regarding the chances of success of his endeavour? If we go back to the passages where Hume introduces supposed ideas, we will notice he points out ideas, especially complex and abstract ones, are by nature faint and obscure, in a way that “’tis almost impossible even for the mind, which forms them, to tell exactly their nature and composition” (Abs. 7, SBN 648–9). This inherent weakness and obscurity of ideas seriously complicates the methodological use of the copy principle, since if we cannot determine the precise structure of complex ideas, we are not in a position to exhaustively search for the impressions that caused them, which leads us to the difficulties mentioned above regarding adequate ideas. At the semantic level, it implies that the meaning of the terms connected to these ideas cannot be fixed with the utmost precision. In consequence, if the meaning of a word depended on the idea attached to it being adequate, then, strictly speaking, there would not be many meaningful words. Besides, in the passages where Hume talks about imperfect ideas, he does not regard the terms connected to them as completely meaningless, but as terms that do not have a determinate meaning.<sup>13</sup> This meaning, although itself imperfect, guarantees a minimum of intelligibility that suffices to apply the copy principle to the imperfect idea they name and proceed with its clarification, at least, up to the point of recognizing the nature of its imperfection.

The faint and obscure nature of ideas gives rise to another difficulty. At EHU 2.9, SBN 21–2 Hume states that these features lead us to confuse ideas that are similar and, consequently, to use terms equivocally, believing that we name one idea when in fact we are naming another. This happens because usually, when we employ a term; we do not have in mind all the features of the idea it names, which would allow us to verify if it is indeed the same idea or a similar one. Thus, Hume observes that “we do not annex distinct and complete ideas to every term we make use of and that in talking of government, church, negotiation, conquest, we seldom spread out in our minds all the simple ideas, of which these complex ones are compos’d” (T 1.1.7.14, SBN 23).

---

13 The only statement of something „completely“ unintelligible occurs when Hume considers the notion of substance of ancient philosophy (T 1.4.3). However, in this case, he seems to point out that the very idea of substance was conceived as „something unknown and invisible, which it supposes to continue the same under all these variations“ (T 1.4.3.4, SBN 220). This suggests that unintelligibility was an inherent quality belonging to the original notion of substance and not a problem pointed out by Hume.

This leads us to the second question, regarding the nature of the connection between ideas and words. Hume explains that terms are usually confusing and ambiguous due to the fact that “we use words for ideas, because they are commonly so closely connected, that the mind easily mistakes them” (T 1.2.5.21, SBN 61–2). Here, Hume describes the relation between ideas and words as a close connection, although he does not further specify the nature of this connection. On the other hand, we have just seen he also points out that sometimes we must „fix“ the meaning of the terms. In the same vein, when referring to the disputes caused by certain ambiguous terms, he indicates that „the disputants affix different ideas to the terms employed in the controversy” (EHU 8.1, SBN 80–1). These clues reveal that the close connection between ideas and words is “common”, that is, general, but not always stable, since in some cases it needs to be properly fixed.

Walter Ott (2006, 236; 2009, 205) suggests that Hume’s theory of linguistic meaning can be understood in the light of the notion of sign he employs when explaining the principle of sympathy. This entails that words are signs that reliably indicate the speaker has a certain idea in mind: “the words or discourses of others have an intimate connection with certain ideas in their mind” (T 1.3.9.12, SBN 112–3). Also, it grants that exposing ourselves to the words of others causes similar ideas in us: “because such a particular idea is commonly annex’d to such a particular word, nothing is required but the hearing of that word to produce the corresponding idea” (T 1.3.6.14, SBN 93) (Ott 2006, 241). Ott proposes, then, that we can suitably understand the connection between ideas and words as a type of causal relation (2006, 236). Moreover, this is suggested by Hume himself in the famous example of the murder of Caesar (T 1.3 4.2, SBN 83). What is the outcome of this interpretation? On the one hand, that there is a general correlation between ideas and words, a correlation that, like all causal relations, is regular but not necessary. On the other hand, the non-necessary character of the relation shows whence problems related to the ambiguity of terms arise: we usually join certain terms with certain ideas, but nothing prevents, in principle, that different connections can take place, or that the words we join to certain ideas be necessarily connected in the same way in other people’s minds.

Now, understanding the connection between ideas and words in terms of a causal relation does allow for the existence of empty terms? Strictly speaking, nothing would prevent us from thinking about this possibility, since it is not contradictory. However, what I want to stress is that the semantic problems Hume deals with do not originate in the fact that the terms in question are literally empty. For example, when he says that it is necessary to „fix“ the meaning of a term, he is not implying the term lacks an idea attached to it, but that it is connected to an imperfect idea. The same happens when he refers to verbal disputes. Hume does not suggest that these disputes originate because the terms are empty, but because the different participants attach different ideas to them. Thus, we can conclude that there are words with more stable

meanings, for a number of reasons: they have a regular relationship with certain ideas, the ideas connected to them are simple, or point to recognizable matters of fact, etc. In sum, they are connected to an adequate idea. But there are other words whose meanings are more fleeting or unstable, because the ideas to which they are connected to are obscure, weak or confusing due to the reasons we have been discussing so far: they are abstract or general, such as „government“ or „negotiation;“ they are extremely complex, like the idea of Paris or a mite; they concern moral issues such as Hannibal’s greatness; they refer to issues exceeding our cognitive capacity, such as the divine mind, etc.

## Conclusion

I hope to have shown that, in the realm of ideas, things are not as simple as the distinction between genuine ideas and fictions suggests. Hume pays close attention to both regular cases and anomalous situations, which is evidenced by the fact that the latter are considered as much —if not more— than the former. An in-depth exam of both instances has revealed many details concerning relations between ideas and their objects. On the one hand, we have seen that ideas that can be regarded as adequate do not always fully capture their objects, but generally do so in a limited way. On the other hand, we have learned that imperfect ideas do not necessarily lead us to error, but rather produce imperfect and unfinished knowledge, and, at the semantic level, give rise to an ambiguity that is the source of disputes and misunderstandings. These shortcomings can be assessed but can only be fixed up to a certain point. In any way, both tasks are relevant: wheter revealing the workings of our imagination, or showing what are the limits of our cognitive skills.

## Bibliography

- Ainslie, Donald (2010). “Adequate Ideas and Modest Skepticism in Hume’s Metaphysics of Space.” *Archiv für Geschichte der Philosophie* 92, pp. 49–51.
- , (2015). *Hume’s True Skepticism*. Oxford: Oxford University Press.
- Costelloe, Timothy M. (2018). *The Imagination in Hume’s Philosophy: The Canvass of the Mind*. Edinburgh: Edinburgh University Press.
- Coventry, Angela (2006). *Hume’s Theory of Causation: A Quasi-Realist Interpretation*. London: Continuum.
- Garrett, Don (1997). *Cognition and Commitment in Hume’s Philosophy*. Oxford: Oxford University Press.
- Garrett, Don (2006). “Hume’s Naturalistic Theory of Representation.” *Synthese* 152, pp. 301–319.

- Hume, David (1739–40/2007). *A Treatise of Human Nature*, edited by David Fate Norton and Mary J. Norton [T]. Oxford: Oxford University Press (1<sup>st</sup> ed., 2000). Cited by book, part, section, paragraph and page number in L. A. Selby Bigge's edition revised by P. H. Nidditch.
- , (1740/2007). “An Abstract of a Book Lately Published: Entituled, *A Treatise of Human Nature*, etc. Wherein the Chief Argument of that Book is farther Illustrated and Explained” [Abs.] In *A Treatise of Human Nature*, edited by David Fate Norton and Mary J. Norton. Oxford: Oxford University Press (1<sup>st</sup> ed., 2000). Cited by paragraph and page number in L. A. Selby Bigge's edition revised by P. H. Nidditch.
- , (1748/1999). *An Enquiry Concerning Human Understanding*, edited by Tom L. Beauchamp [EHU]. Oxford: Oxford University Press. Cited by section, paragraph and page number in L. A. Selby Bigge's edition revised by P. H. Nidditch.
- , (1779/1947). *Dialogues Concerning Natural Religion*, edited with an introduction by Norman Kemp Smith [DNR]. Indianapolis: Bobbs-Merrill. Cited by part, paragraph and page number in Kemp Smith edition.
- Kail, Peter (2003). “Conceivability and Modality in Hume: A Lemma in an Argument in Defense of Skeptical Realism.” *Hume Studies* 29 (1), pp. 43–61.
- Landy, David (2018). *Hume's Science of Human Nature. Scientific Realism, Reason, and Substantial Explanation*. New York: Routledge.
- Lightner, D. Tycerum (1997). “Hume on Conceivability and Inconceivability.” *Hume Studies* 23 (1), pp. 113–132.
- Locke, John (1689/1975). *An Essay Concerning Human Understanding*, ed. P. H. Nidditch [E]. Oxford, Clarendon Press. Cited by book, chapter and section.
- Ott, Walter (2006). “Hume on Meaning.” *Hume Studies* 32 (2), pp. 133–252.
- , (2009). *Causation and Laws of Nature in Early Modern Philosophy*. Oxford: Oxford University Press.
- Powell, Lewis (2013). “How to Avoid Mis-Reading Hume's Maxim of Conceivability.” *The Philosophical Quarterly* 63 (250), pp. 105–119.
- Richman, Kenneth N. (2000). “Introduction,” in Read, Rupert and Richman, Kenneth N. *The New Hume Debate*. London: Routledge, pp. 1–15.
- Traiger, Saul (1987). “Impressions, Ideas, and Fictions.” *Hume Studies* 23 (2), pp. 381–399.



Gabriel Watts  
Centre for Advanced Studies  
in the Humanities – Human Abilities,  
Freie Universität Berlin  
gabriel.watts@sydney.edu.au

Original Scientific Paper  
UDC 165 Xjym Д.  
165.195

## READING HUME ON THE PASSIONS

**Abstract:** This paper provides a reception history of Book Two of the *Treatise – Of the passions* – as well as an attempt to reconcile Hume’s ambitions to systematicity in Book Two with the distracted and distracting nature of the text. We currently have, I think, a good sense of the philosophical importance of Book Two within Hume’s science of human nature. Yet we have not made much progress on understanding Book Two on its own terms, and especially why Book Two so often seems on the verge of falling into an explanatory heap. I aim to rectify this situation by giving a reading of Book Two that makes sense of the philosophical importance of Hume’s system of the passions, yet also explains why he encounters so many difficulties in setting out his system; such that he is often forced to stretch his explanations to the very edge of the credible. I contend that Hume’s system of the passions is best viewed as an unstable explanatory compound, one that progressively dissolves as Hume’s explanatory intentions become increasingly ambitious.

**Keywords:** Hume, Passions, Emotions, *Treatise*, Animals

### 1. Introduction

Of the three books that comprise Hume’s *A Treatise of Human Nature*, it is Book Two, ‘*Of the passions*’, that most reads as if it was written by a man in his mid-twenties. At one point Hume says that ‘[t]he different stations of life influence the whole fabric [of a person], internal and external’ (T 2.3.1.9, SBN 402) and it is difficult not to apply this to Hume’s own concerns, intellectual and otherwise. He is very interested to understand the origin of our desire for riches and power (T 2.1.10, 2.2.5; SBN 309–316, 357–365), and he is acutely sensitive to the shame felt by those of ‘narrow circumstances’ (although not especially empathetic) (T 2.1.11.14, SBN 322). Sexually-charged relationships are frequently mined for examples (T 2.1.11.19, 2.2.11, 2.3.4.3; SBN 324, 394–6, 420), as well as other ‘curious phenomena’ such as the weakening of the maternal bond after a second marriage (T 2.2.4.9, SBN 355). And this is without even mentioning the many inconsistencies – both terminological and substantive – that arise from the tension between the young Hume’s excessive desire to explain and his commitment to producing a system of the passions that is ‘so simple’ as to compel his reader’s assent (T 2.1.12.9,

2.2.12.1; SBN 327–8, 397). Indeed Book Two is so stuffed with examples, often requiring searching and unwieldy explanations, that it is perhaps a minor miracle that Hume's philosophy of the passions has come to be seen as especially important for understanding his intentions in the *Treatise*, if not his philosophical intentions on the whole.

In this paper I take up the question of the explanatory unity of Book Two of the *Treatise*; or rather, its seeming lack thereof. I begin by giving a reception history of Book Two. My aim here is to show that Hume's system of the passions was initially regarded as a distracted and distracting text with no strong connection to the rest of his philosophy, and that our current sense of Book Two's philosophical importance has largely been shaped in response to this false charge: either by reading Book Two in light of Hume's concerns in Book One, Book Three, or both; or in light of broader philosophical concerns regarding the government of the passions. However, given that we now have a well-developed sense of the value and importance of Book Two within these contexts, I suggest that we examine Book Two on its own terms, and see if we cannot find unity of purpose amongst Hume's intentions for his philosophy of the passions, considered as an important part of his science of man in its own right.

My claims are schematic, but what I suggest is that the philosophy of the passions that Hume sets out in Book Two is shaped by three distinct intentions. Primarily, Hume's philosophy of the passions is an attempt to explain why increasing affluence in any society necessitates the establishment of civil government. Most of Hume's key analyses and innovations in Book Two, I claim, are well-understood as having been originally invented to this purpose. Secondly, Hume's philosophy of the passions is an attempt to draw sense-based theories of mankind's natural benevolence within the framework that he had invented to account for the origin of civil government. However, as it turns out, the system of the passions that Hume had invented to explain the origin of civil government is not especially well-suited to this purpose, and most of the explanatory difficulties that Hume encounters in Book Two can be understood in light of his attempt to stretch his philosophy of the passions beyond its original purpose.

I then suggest that these primary and secondary intentions are subsumed under a third, which gives Hume's philosophy of the passions its unity of purpose. Namely, Hume's philosophy of the passions is an attempt to explain human sociability using principles that are observed to operate throughout the whole of 'sensitive creation.' This was part of Hume's design in first developing his system of the passions, and it motivates him to explain our natural benevolence using the same system. However, Hume's prosecuting this plan across the whole of Book Two requires him to explain the origin of curiosity – a sociable passion that we do not share with animals – using principles that are common to both man and beast. Hume is able to do so, I claim, but only at the cost of obscuring his initial distinction between the 'direct' and 'indirect' passions.

## 2. The Reception of Hume's Philosophy of the Passions

A good place to begin a reception history of Hume's philosophy of the passions is with Hume's own reworking of Book Two in *A Dissertation on the Passions*, which he published as part of his *Four Dissertations* (1757). The *Dissertation* is a truncated version of the text of Book Two of the *Treatise*: it is notably shorn of examples, has some structural changes, and seems to make a few conceptual adjustments.<sup>14</sup> With its final paragraph, Hume sums up his aims for the *Dissertation* with the remark that:

I pretend not to have here exhausted this subject. It is sufficient for my purpose, if I have made it appear, that, in the production and conduct of the passions, there is a certain regular mechanism, which is susceptible of as accurate a disquisition, as the laws of motion, optics, hydrostatics, or any part of natural philosophy.

(DP 6.19)

While we ought not to suppose that Hume's overall purpose with the *Dissertation* was the same as with Book Two of the *Treatise*, I think it is safe to assume that Hume intended both tracts to show that the 'production' and 'conduct' of the passions can be traced to 'regular' mental mechanisms, and that these are 'susceptible of as accurate a disquisition, as the laws of motion, optics, hydrostatics, or any part of natural philosophy'.

Indeed that this was amongst Hume's aims with Book Two has never really been in doubt. What has been vigorously debated, however, is the philosophical point, and value, of Hume's attempting to do so. As regards his attempt to show that the 'conduct' of our passions is not the product of a faculty of practical 'reason', but is entirely reducible to mechanical principles, Hume has been severely chastened; both for his execution and his intentions. Thomas Reid regarded Hume's denial of rational principles of action – encapsulated in Hume's infamous dictum that 'reason is, and ought only to be the slave of the passions' (T 2.3.3.4, SBN 414–5) – as hopelessly confused,<sup>15</sup> a judgement that many others have shared.<sup>16</sup> T.H. Green goes even further. In his general introduction to Books Two and Three of the *Treatise*, Green

---

14 For the only extended study of the *A Dissertation on the Passions* in relation to Book Two of the *Treatise* (that I know of), see Amyas Merivale, *Hume on Art, Emotion, and Superstition* (New York: Routledge, 2018), esp. 128–44; for a detailed account of the textual truncations, see Tom Beauchamp, "Introduction: A History of Two Dissertations" in David Hume, *A Dissertation on the Passions: The Natural History of Religion: A Critical Edition*, ed. Tom Beauchamp (Oxford: Oxford University Press, 2007), l-cxvi.

15 Thomas Reid, *Essays on the Active Powers of Man*, eds. Knud Haakonssen and James A. Harris (Edinburgh: Edinburgh University Press, 2010), 157; see also 135–6, 153.

16 Even those who are otherwise favourably disposed towards Hume's moral theory have been relatively unimpressed. See here, J.L. Mackie, *Hume's Moral Theory* (London: Routledge and Keagan Paul, 1980), 44–63.

characterises Hume's mechanistic account of volitional actions as a deliberate attempt to 'degrade' morality by pouring 'solvent' on reason.<sup>17</sup>

Green is similarly dismissive of Hume's account of the causal origins of various passions, the exposition of which makes up the bulk of Book Two.<sup>18</sup> Green presents Book Two – the very work he is introducing – as proceeding from false assumptions, and, as a result, as being full of explanatory 'subterfuges'<sup>19</sup>. An especially entertaining example concerns Hume's attempt to explain the 'indirect' origin of pride and humility, as well as love and hatred, from what he (Hume) calls a 'double relation of ideas and impressions'<sup>20</sup>. Here Green wryly observes that to make his explanation work, Hume introduces a principle of association between impressions which is similar to, but slightly different from, the principle of association between ideas that he had introduced in Book One. This is 'somewhat strange', Green says, given that Hume had previously claimed that there was no essential difference between impressions and ideas, but rather that ideas were just phenomenally weaker 'copies' of impressions:<sup>21</sup> for, if this were the case, ought any principle of association between ideas apply equally to impressions? Indeed, almost any objection that might be levelled at the coherence of Hume's philosophy of the passions can be found in Green's introduction; the writing of which Green described as an 'irksome labour' whose chief goal was to show that there was nothing of value in Hume's 'anachronistic system'<sup>22</sup>.

Interestingly, Green's view of Hume's philosophy of the passions as containing nothing much of value was endorsed by Norman Kemp Smith, who otherwise set himself in direct opposition to the Reid-Green interpretation of Hume as a sceptical philosopher.<sup>23</sup> On the Reid-Green reading, as Kemp Smith describes it, 'Hume's teaching is sheerly negative, being in effect little more than a *reductio ad absurdum* of the principles which Hume's predecessors, and Hume himself, have followed in their enquiries'<sup>24</sup>. To those who accepted this reading of Hume's philosophy, Hume's scepticism in morals was of a piece with his scepticism in epistemology, and his overall

17 T. H. Green, "Introduction to Moral Part of the *Treatise*" in *The Philosophical Works of David Hume*, 4 vols., eds. T. H. Green and T. H. Grose (London: Longmans, Green, and Co, 1874–5), vol 2 (1874), 1, 48–54.

18 Green, 32–48.

19 Green, 37.

20 See here, Green, 35–42.

21 Green, 36.

22 Green, 71.

23 See Norman Kemp Smith, *The Philosophy of David Hume* (New York: Palgrave, Macmillan, 1941) (*The Philosophy of David Hume*), 3–14; Norman Kemp Smith, "The Naturalism of David Hume (I)", *Mind* 14.54 (April 1905), 149–73; and Norman Kemp Smith, "The Naturalism of David Hume (II)", *Mind* 14.55 (July 1905), 335–47. See also on this point, John Laird, *Hume's Philosophy of Human Nature* (London: Routledge, 1932).

24 Kemp Smith, *The Philosophy of David Hume*, 3.

philosophy was a potent intellectual solvent, to be handled very carefully, and only as necessity required.

Against this, Kemp Smith claimed that the ‘correct reading’ of Hume’s purposes with his sceptical arguments was to show that our most fundamental beliefs – our belief in the existence of external objects, in our personal identity, and in the rational warrant of our arguments – are the products of ‘feeling, not reason or understanding ... and therefore also not ideas – at least not ‘ideas’ as hitherto understood<sup>25</sup>. Hume’s most ‘general title’ for such ‘feelings’, Kemp Smith claimed, is “passion’ ... and belief, he teaches, is a passion<sup>26</sup>. This being the case, Kemp Smith asserted that:

The maxim which is central in his ethics – ‘Reason is and ought to be the slave of the passions’ – is no less central in his theory of knowledge, being there the maxim: ‘Reason is and ought to be subordinate to our natural beliefs.’<sup>27</sup>

It was this general disposition of thought that Kemp Smith coined Hume’s ‘naturalism’, and he claimed that it represented Hume’s fundamental philosophical disposition, against which his sceptical arguments, as well as everything else in his philosophy, ought to be understood.

Kemp Smith’s conception of Hume as a fundamentally naturalistic philosopher has been widely influential. As such, it is noteworthy that Kemp Smith himself had a low opinion of Book Two. In a section of *The Philosophy of David Hume* with the amusing title, ‘*Unsatisfactory Features of the Argument and Exposition in Book II*’, Kemp Smith writes (and I quote at length):

For several reasons Book II, as regards sequence and mode of exposition, is the least satisfactory of the three books which constitute the *Treatise* ... More than a third of Book II is employed in the treatment of four passions which have no very direct bearing upon Hume’s ethical problems, and play indeed no really distinctive part in Hume’s system – pride, humility, love and hatred ... In so far as Hume’s purpose in discussing these four passions is to support his thesis that the laws of association play a rôle in the mental world no less important than that of gravity in the physical world, his argument does connect itself with Book I ... But even so, he bewilders his readers by introducing two special laws of association (one of them between *impressions!*) additional to those mentioned in Book I, and by an over-ingenuous elaboration of his argument. This is, indeed, the most outstanding instance in which Hume’s secondary plot – a statics and dynamics of the mind – has broken in upon, and has unhappily thrown into confusion, the requirements proper to his main programme.<sup>28</sup>

---

25 Kemp Smith, *The Philosophy of David Hume*, 11.

26 Kemp Smith, *The Philosophy of David Hume*, 11.

27 Kemp Smith, *The Philosophy of David Hume*, 11.

28 Kemp Smith, *The Philosophy of David Hume*, 160–1.

This passage is worth quoting at length because it reveals that Kemp Smith's disagreement with Green and others as to the merits of Hume's philosophy was a disagreement over the merits of Hume's epistemology. Regarding the philosophy of the passions that Hume develops in Book Two, Kemp Smith mostly agreed that it was a 'hindrance'<sup>29</sup>.

The passage above is also useful insofar as Kemp Smith's criticisms of Book Two served as grist to the mill for a later generation of scholars, who – while accepting Kemp Smith's thesis that Hume's fundamental philosophical disposition was naturalistic, rather than sceptical – set out to prove that Kemp Smith had underestimated the importance of Hume's philosophy of the passions to his overall philosophy of human nature. Here Páll Árdal took it upon himself to demonstrate the crucial importance of the passions of pride, humility, love and hatred to Hume's moral philosophy – in particular their close connection to moral sentiments.<sup>30</sup> While Nicholas Capaldi sought to show that Hume's 'statics and dynamics of the mind' was not a 'secondary plot', as Kemp Smith had claimed, but was the *primary* plot, and that Hume's causal account of the origins of our passions played a leading role.<sup>31</sup>

More precisely, Capaldi argued that Kemp Smith's thesis regarding 'the primacy of feeling in Hume' had the 'unfortunate' effect of making Hume's philosophy 'even more irrationalist than in the Reid interpretation'<sup>32</sup>. Kemp Smith had made this error, Capaldi thought, because he had failed to see 'the real connection of psychology and logic in Hume'<sup>33</sup>. However, once one came to see that Hume's epistemological naturalism was grounded in his 'Newtonian' thesis regarding the role that 'laws of association' play in the mental world, then one would come to see that Hume's theory that our passions arise from various principles of mental association is an essential part of Hume's science of human nature, not a hindrance for Hume's readers to overcome.<sup>34</sup> Indeed, by 1975, Capaldi could write that:

Book II of the *Treatise* is in many respects the most important for exemplifying the major themes of Hume's philosophy. By its discussion of the mechanism of association it serves as a confirmation of Hume's explanation of the mechanics of causal belief. At the same time, it prepares the way for Hume's moral theory by explaining the

---

29 Kemp Smith, *The Philosophy of David Hume*, 161.

30 See esp. Páll Árdal, *Passion and Value in Hume's Treatise* (Edinburgh University Press, 1966), 109–47.

31 Nicholas Capaldi, *David Hume, The Newtonian Philosopher* (Boston: Twayne Publishers, 1975), 130–50.

32 Capaldi, 229.

33 Capaldi, 229.

34 See here, Capaldi, 49–70. In arguing like this Capaldi can be seen to be following Laird, who was the first to stress the significance of Newton for Hume, or that 'Hume meant, in short, to become the Newton of the Human Mind': Laird, 29.

mechanism of sympathy. Again, it serves as an explanation for the concept of the self which proved so problematic in Book I. Finally, by outlining his theory of human motivation Hume completes his project of undermining both the rationalist model in general and rationalist moral theories. Thus, an understanding of the passions is essential for an understanding of all the philosophical issues in Hume's *Treatise*.<sup>35</sup>

This claim, now forty-five years old, gives a good general description of the scholarly outlook that has characterised the study of Book Two ever since; for the details of which I refer the reader to Elizabeth Radcliffe's excellent overview of the current state of the art in her article 'Hume's *Psychology of the Passions*' (2015).<sup>36</sup> Instead I want to focus on one particular line of argument that has come to the fore as a means of assessing, and asserting, the importance of Hume's philosophy of the passions as a distinctive concern within Hume's philosophy of human nature. This line of argument is nicely foreshadowed by Capaldi's claim that with Book Two 'Hume completes his project of undermining both the rationalist model in general and rationalist moral theories'; yet its first sustained articulation is given in Annette Baier's book *A Progress of Sentiments* (1991).<sup>37</sup>

Baier's fundamental assertion is that Books One and Two of the *Treatise* are linked together by a narrative dimension. Baier's core claim – in her own words – is that with his depiction of his skeptical crisis at the end of Book One, Hume:

enacts for us the turn he wants us to imitate, a turn from a one-sided reliance on intellect and its methods of proceeding to an attempt to use, in our philosophy, *all* the capacities of the human mind: memory, passion and sentiment as well as a chastened intellect. That is what Hume attempts from Book Two onwards – not only are passions his topic there, but his approach to them is to be guided by experience-informed passion, and he recommends that his readers indulge their sentiments when they join him in his pursuit of philosophy “in this careless manner”.<sup>38</sup>

On Baier's view, Hume's *Treatise* is something of a three-act epistemological drama, in that, '[l]ike Descartes in his *Meditations*, so Hume in his *Treatise* stages a thinker's dramatic development from inadequate and doubt-inviting

---

35 Capaldi, 130.

36 Elizabeth S. Radcliffe, "Hume's Psychology of the Passions: The Literature and Future Directions," *Journal of the History of Philosophy*, 53:4 (2015): 565–605.

37 For similar ideas, articulated at a similar time see, Amélie Oksenberg Rorty, "Pride produces the idea of self": Hume on moral agency," *Australasian Journal of Philosophy*, 68:3 (1990): 255–269.

38 Annette Baier, *A Progress of Sentiments: Reflections on Hume's Treatise* (Cambridge, MA: Harvard University Press, 1991), 1.

approaches to more satisfactory reflections'<sup>39</sup>. Hume's genius, however, was to have shown that what Baier calls the 'Cartesian' conception of rationality is self-undermining, and that the recognition of this ought to push us towards an 'embodied', 'social' conception of ourselves as rational beings: which is to say, towards a philosophical self-conception that we, and others, endorse upon reflection, and, as such, is capable of withstanding skeptical challenges of the type that Hume encounters at the end of Book One.<sup>40</sup>

Baier herself could be rather careless with the details of Hume's philosophy. Yet she gives a compelling account of the general intellectual thrust of the *Treatise*, and a number of later scholars have stepped in to fill out the specifics.<sup>41</sup> The most notable adjustment of Baier's basic position, in my view, is the recent turn towards understanding the relationship between Books One and Two of the *Treatise* in light of the broader philosophical concern with the relation between reason and passion, rather than through Baier's more narrowly epistemological framing. James Harris puts the point succinctly when he says that:

Missing from Baier's account, however, is full acknowledgement of the magnitude of the problem that Hume has created for himself as he moves on to Book Two ... What Hume knew, and what his readers knew, was that the passions were usually taken to be a problem that required reason for its solution. How, without governance by reason, could the passions supply anything other than distraction and disturbance?<sup>42</sup>

Broadly speaking, answers to this question have gone two ways. The answer that Harris, and others, propose is that with Book Two of the *Treatise* Hume intends show that the passions are capable of governing themselves.<sup>43</sup> More

---

39 Baier, vii.

40 For Baier's final articulation of her position see Annette Baier, "Reflexivity and Sentiment in Hume's Philosophy", in *The Oxford Handbook of Hume*, ed. Paul Russell (Oxford: Oxford University Press, 2016), 54–9.

41 See here: Lilli Alanen, "Reflection and Ideas in Hume's Account of the Passions", in *Persons and Passions: Essays in Honour of Annette Baier*, eds. Joyce Jenkins, Jennifer Whiting, and Christopher Williams (University of Notre Dame Press, 2005) (*Persons and Passions*), 117–43; Donald Ainslie, "Sympathy and the Unity of Hume's Idea of Self", in *Persons and Passions*, 143–73; Janet Broughton, "Hume's Voyage" in *Persons and Passions*, 174–91; for book-length studies in this vein, see Jacqueline Taylor, *Reflecting Subjects: Passion, Sympathy, and Society in Hume's Philosophy*, (Oxford: Oxford University Press, 2015), and Jay L. Garfield, *The Concealed Influence of Custom: Hume's Treatise from the Inside Out*, (New York: Oxford University Press, 2019).

42 James A. Harris, "'A Compleat Chain of Reasoning'; Hume's Project in A Treatise of Human Nature, Books 1 and 2", *Proceedings of the Aristotelian Society*, 109 (2009), (*Compleat Chain*) 135.

43 See here, Jane L. McIntyre, "Hume's 'New and extraordinary' account of the passions", *The Blackwell Guide to Hume's Treatise*, ed. Saul Traiger, (Malden, MA: Blackwell, 2006), 199–215; Christopher J. Finlay, *Hume's Social Philosophy: Human Nature and Commercial Sociability in A Treatise of Human Nature* (Continuum, 2007), 44–157; Taylor, 1–98; see



precisely, the claim is that Hume deliberately analyses passions that had traditionally been portrayed as socially destructive forces – pride, vanity, the desire for riches and power – and shows them to be mechanically tuned towards pro-social ends on account of our sympathetic ‘concern for how others see us, a desire for their love and admiration, [and] an aversion to their hatred and contempt’<sup>44</sup>. In this way, Hume is thought to establish the existence of a self-ordering realm of passionate activity. One that is independent, in principle, from the political realm; such that the government of the passions is revealed as being to a considerable extent *self-government* – ‘in the context of the social realm taken as a whole’<sup>45</sup> – rather than political government.

Mikko Tolonen, by contrast, has argued that Hume’s basic position on the control of the passions is that there is no ‘lasting spontaneous order in a large society to serve self-interest or other passions without a government to redirect their natural course’<sup>46</sup>. As such, whatever capacity for self-government the passions might have, it only operates small rudimentary societies, not large civil societies ‘in which people are not necessarily related to or even acquainted with each other’<sup>47</sup>. It is Tolonen’s view that the principle of sympathy ‘is not sufficient to overcome the challenges that passions create in large societies’<sup>48</sup>: a point that he thinks Hume makes abundantly clear in Book Three of the *Treatise*. Thus when it comes to the question of government of the passions, we find that our ‘rechanneling the passions through political and moral customs still plays the key role’<sup>49</sup>. It is Tolonen’s contention that Hume’s focus on our self-liking and self-interested passions in Book Two was intended to set up his Book Three view that the origins of our most important political and moral conventions can be traced to the naturally disruptive consequences of pride, vanity, and self-interest for large-scale collective living.

It is not my design to adjudicate between these interpretations here. Other than to say that the social world of Book Two is one in which the conventions that establish the ranks of civil authority are already in a highly developed state. This makes it very difficult to tell whether Hume’s presentation of the passions as forming a sympathetically self-regulating economy, governed primarily by considerations of honour and dishonour, respect and contempt,

---

also James A. Harris, *Hume: An Intellectual Biography* (Cambridge: Cambridge University Press, 2015), 102–116.

44 Harris, *Compleat Chain*, 137.

45 Harris, *Compleat Chain*, 132.

46 Mikko Tolonen, *Mandeville and Hume: Anatomists of Civil Society* (Voltaire Foundation: Oxford, 2013), 13–4.

47 Tolonen, 13.

48 Tolonen, 22.

49 Tolonen, 22.

in fact merely reflects the long-established conventions of civil government and polite behaviour that keep the various 'ranks' of individuals in steady place. Yet, whatever the case, there is general agreement that Hume dismisses the notion that a distinct rational principle of authority, such as conscience or practical reason, is required for control of the passions; and that whatever the aims of Hume's theory of the 'indirect' passions were, Hume makes a novel contribution to eighteenth century debates about human sociability.

What this reception history of Book Two makes clear, I hope, is that the last seventy-odd years of scholarship have established the importance and value of Hume's philosophy of the passions, both within his *Treatise* project and in the context of larger debates over the relationship between reason and the passions. Yet, by the same measure, I think it is safe to say that our current understanding of Hume's intentions with Book Two has been almost entirely imposed from the outside. We have sought to understand Hume's philosophy of the passions in the context of the *Treatise* as a whole, or as first published, or as part of larger philosophical discussions about the nature of the passions. Not so much on its own terms.

There are perfectly good reasons for this. As we saw, the study of Book Two as an important part of Hume's science of man was initially a response to early assessments of the low philosophical value of Book Two, both from Hume's detractors and his chief defenders, and not only within his *Treatise* project, but also concerning the general question of how the passions ought to be controlled. As a result of this scholarship, and that which has since built upon it, we have developed a good sense – a fine-grained sense – of what Hume was up to in Book Two. Yet we have not, I think, made much progress on the question of the internal explanatory unity of Hume's philosophy of the passions; or rather, on the question of why Hume's attempt to systematise human feeling and action lacks a clear unity of purpose.

Book Two does, of course, have some notable structuring features, namely Hume's novel division of the passions into those that are 'direct' and those that are 'indirect'. But even this is not a division that Hume makes, or keeps, especially clear. Then there are questions about his inclusion of the will within his treatment of the passions, even though 'properly speaking' it does not belong there (T 2.3.1.2, SBN 399), including the question of whether or not he thinks that it is, in fact, a passion (see T 2.3.9.7, SBN 439). There is also his decision to seemingly tack an account of curiosity onto the end of Book Two, because it would have been 'impossible' to have included it 'under any other head ... without danger of obscurity and confusion' (T 2.3.10.1, SBN 448). And what to make of his introduction of 'blended' passions in the middle of his account of love and hatred (T 2.2.6.1, SBN 366)? This certainly seems to be an assertion of a categorical distinction between impressions and ideas; flatly contradicting his earlier claims that these two classes of perceptions are 'essentially the same' and distinguished

only by their ‘liveliness and vivacity’<sup>50</sup>. Jane McIntyre has referred to such phenomena as giving an ‘elusive’ quality to Book Two, which is a polite way of putting it.<sup>51</sup> Much more than the other books of the *Treatise*, if not the rest of Hume’s philosophical writings, Book Two gives the appearance of having been worked out as he went along: Reid, Green, and Kemp Smith make a good point when they portray Book Two as both distracted and distracting when read on its own terms.

### 3. The Internal Unity of Hume’s Philosophy of the Passions

Still, now that we have a much better sense of the value and importance of Book Two, I want to suggest that we focus our attention on the question of the internal unity of Hume’s philosophy of the passions. This is not to say, to be clear, that we should foreground issues of textual exegesis regarding Hume’s key distinctions: between direct and indirect passions, passions and sentiments (and emotions), or between calm and violent passions. This is important work, but I doubt that we will ever come to grips with what Hume “really means” when he draws such distinctions, given how careless he is with his terminology;<sup>52</sup> and anyway, to the extent that we might look for unity amongst any of Hume’s philosophical writings, our model ought not to be that of a textual puzzle awaiting solution: “Hume’s true meaning”. Surely Hume himself would not have wanted us to treat his philosophy as a static object. Instead our model should be that of a river. Something fundamentally in motion, but which nevertheless has consistent general features. It is these general points of consistency that I am interested in.

Nor do I mean to say that we ought to study Hume’s philosophy of the passions without any reference to his intellectual context, or to his other philosophical writings, and especially to the other two books of the *Treatise*. This would be to embrace a bizarre sort of interpretive asceticism. Rather, my suggestion is that when we do look to external context, or to Hume’s other writings, it is in the service of reading Book Two on its own terms, as compared, say, to using Book Two to either confirm or complete points made in Book One, or as serving as the foundation of the moral theory that he expounds in Book Three.<sup>53</sup>

50 For examination of this point within Hume’s broader explanatory framework, see Tamás Demeter, *David Hume and the Culture of Scottish Newtonianism Methodology and Ideology in Enlightenment Inquiry* (Boston: Brill, 2016), 134–149.

51 McIntyre, 199.

52 This is not to say that Hume’s carelessness is not relatively consistent, or sometimes merely a matter of stylistic variation, such that we cannot get a good idea of what he was thinking. But rather that his carelessness leaves enough room to doubt that we could ever settle such questions authoritatively.

53 Hume does say that his philosophy of the passions serves as the ‘foundation’ of the rest of his science of human nature (T Abstract 3, SBN 646). My point is simply that we are not bound to reading it as such.

For instance, as Harris, Tolonen, and others have shown, a central concern of Hume's philosophy of the passions is the question of the government of the passions, and that much of what Hume says in Book Two can be seen as setting up, or serving as the psychological foundation for, claims that he wants to make about the origin of government in Book Three; especially as regards mankind's natural capacity to live *without* government for considerable stretches of time, at least in certain circumstances (see T 3.2.8.1–3, SBN 539–42). Yet we might also see Hume's genealogy of government as being empirically well-founded, independent of his psychology of the passions; given that we have knowledge of (what Hume thinks are) rudimentary societies that exist without civil government, but can also plainly see that most societies establish some kind of sovereign political authority as they grow in size and stature (T 3.2.8.1–3, SBN 539–42). On this reading, Hume's psychology of the passions is an independent (although closely related) attempt to understand the psychological reasons behind the genealogical observation that government becomes 'absolutely necessary' once certain conditions are met (see T 3.2.8.1, SBN 539); and more precisely, why it is that '[n]othing but an encrease of riches and possessions' can 'oblige' human beings to establish civil magistrates (see T 3.2.8.2, SBN 541).

If we read things this way we can approach Book Two quite differently to how it reads on the page. Here, rather than beginning with Hume's account of pride and humility as phenomenally simple indirect impressions of reflection, we can begin by looking at his explanation of our desire for riches and possessions. In particular, with his claim that 'one of the principal recommendations of riches, and ... the chief reason, why we either desire them for ourselves, or esteem them in others' is the 'secondary satisfaction or vanity' that 'a rich man receives from his possessions' (T 2.2.5.21, SBN 365). This 'secondary satisfaction or vanity' is, of course, an instance of pride. But it is one with an importantly complex aetiology.

For starters, the original feeling of pride that the rich and propertied take in their possessions and riches is itself a rather complicated affair. Such feelings arise, Hume thinks, not simply from their considering the utility of their wealth and property to themselves, or to others, but from their pleasing belief – whether true or false – that they are at 'liberty' to reap whatever advantages they choose, by the use of their riches and property, within the bounds of 'justice and moral equity' (T 2.1.10.1–12, SBN 309–16). Hume's argument for this claim is impossible to unpack here, but it demands, as he acknowledges, an account of the nature of human liberty, and more precisely, of the belief that wealth confers a power to act, or not act, as one pleases (see here T 2.1.10.9, SBN 314).<sup>54</sup>

---

54 For more on these points see Taylor, 71–98.

Further, the ‘love and esteem’ that we have for the rich and powerful is not (for the most part, Hume thinks) a matter of our considering the utility that their possessions and wealth afford them (see here T 2.2.5, SBN 357–65). Rather it is a matter of our sympathising with ‘the satisfaction, which this power [to act or not act as one pleases] affords the person, who is possessed of it’ (T 2.2.5.6, SBN 359). On Hume’s theory that both love and hatred, like pride and humility, are caused by a double relation of ideas and impressions, our sympathetic perception of the satisfaction that the rich and propertied take in their liberty to act on their desires naturally causes us to love them. Our esteem for such persons, insofar as we might distinguish esteem from love, is then a mixture of this love with humility, the latter of which being caused by our comparing their possessions and wealth to our own circumstances and finding ourselves wanting (see here T 2.2.10, SBN 389–93).

The ‘secondary satisfaction or vanity’ that the rich and propertied take in their wealth and possessions is then a further sensation of pride that arises when such persons sympathise with their admirers and perceive the pleasing sentiments of love and esteem that their (own) riches and possessions produce (T 2.2.5.21, SBN 365). Not only this, but in instances where their riches and possessions afford them the power of using other persons as the means to the satisfaction of their desires, this ‘vanity of power’ is augmented by a further comparative pleasure that arises from their comparing the superiority of their power to that of other rational creatures (T 2.1.10.12, SBN 315–6). Nor is this a throwaway point. Rather, Hume contends that:

supposing it possible to frame statues of such an admirable mechanism, that they cou’d move and act in obedience to the will; ‘tis evident the possession of them wou’d give pleasure and pride, but not to such a degree, as the same authority, when exerted over sensible and rational creatures, whose condition, being compar’d to our own, makes it seem more agreeable and honourable. Comparison is in every case a sure method of augmenting our esteem of any thing. A rich man feels the felicity of his condition better by opposing it to that of a beggar.

(T 2.1.10.12, SBN 315–6)

Hume at one point seems to call this particular species of vanity ‘ambition’, and to suggest that this passion is a species of vanity that is augmented by a malicious desire to subjugate other rational creatures so as to experience and maintain a pleasing comparison (T 2.2.8.14, SBN 378). But, as with many other things, he never makes this clear.

Still, it seems clear enough that Hume regarded the desire of the rich and propertied to maintain the power that their riches and possessions afford them, as well as the desire of those who esteem the rich and propertied to acquire such power for themselves, to be the reason why an ‘increase in riches and possessions’ in any society naturally leads to the establishment of

government. For, on the one hand, those who are comparatively poor and possessionless see an opportunity to increase their power, while, on the other, those who are already rich and propertied see an opportunity to further increase their dominion over others. As the riches and possessions on offer are limited, these two interests are in conflict. And since Hume thinks that conflicts between passions naturally tend to obscure our judgements, such that we prefer what is near and at hand to the remote goods that obeying the laws of justice guarantees us (T 3.2.7, SBN 534–9), the establishment of some sort of civil authority becomes ‘absolutely necessary’: both to enforce the laws of property possession, and to decide any disputes over what those laws require (T 3.2.7, SBN 534–9).

It is perhaps not very surprising that Hume sees avarice and ambition at the root of the need for government. More notable, in this regard, is his detached “anatomical” method of analysing these phenomena.<sup>55</sup> My purpose in sketching out Hume’s main moves, however, is to suggest that the complexity of his explanation of the desire for riches and property – an explanation that employs ideas from across the whole of Book Two, and relies heavily upon his key conceptual devices: sympathy, comparison, and double relations of ideas and impressions – can be seen as evidence that his philosophy of the passions was primarily designed to account for this passion; rather than it being fortuitous that Hume’s association-driven account of the mechanisms of our passions just happens to be able to explain why an increase in material wealth necessitates the establishment of political authorities, and to do so in a painfully precise manner.

One way to confirm this suspicion is to pay attention to those points where Hume confesses to find special difficulty in accounting for certain phenomena within his explanatory system. Of particular note here is his observation that ‘[t]he passions of love and hatred are always followed by, or rather conjoin’d with benevolence and anger’ (T 2.2.6.3, SBN 367). The observation itself is plain enough, but, as Hume laments, it does not ‘quadrate exactly with the principles that we would endeavour to establish’ (T 2.2.6.2, SBN 366), namely, his theory that love and hatred originate from a double relation of ideas and impressions, just as pride and humility do. For, insofar as love and hatred do arise in the same manner as pride and humility, Hume is unable to explain why love and hatred, unlike pride and humility, ‘are not completed within themselves, nor rest in that emotion, which they produce, but carry the mind to something farther’ (T 2.2.6.3, SBN 367).

Instead Hume is forced to either (1) concede that ‘love is nothing but the desire of happiness to another person, and hatred that of misery’ (T 2.2.6.4, SBN 367), which would make these passions ‘direct’ passions, on his system; or (2) to assert that love, as he has accounted for it, is conjoined with a separate passion of benevolence, and hatred with a separate passion of anger, ‘by the

55 See here Tolonen, 1–40.

original constitution of the mind' (T 2.2.6.6, SBN 368). Hume chooses (2). Yet, in doing so, he must concede that our benevolence towards those whom we love does not arise from a prior perception of the pleasure that our benevolent actions will cause them, nor does our anger towards those whom we hate arise from an appreciation of the pain that our hateful actions will cause them (T 2.2.6.5–6, SBN 367–8). Instead these passions inexplicably arise whenever we feel love or hatred towards someone. This, however, requires Hume to posit that benevolence and anger are 'direct' passions that are not caused by perceptions of pleasure or pain (T 2.3.9.8, SBN 439); a move which complicates his initial description of 'direct' passions as those that 'arise immediately from good and evil, from pain or pleasure' (T 2.1.1.4, SBN 276).

Elizabeth Radcliffe has argued that Hume's analysis of loving benevolence and hateful anger into sets of distinct impressions is a deliberate rejection of Francis Hutcheson's account of the nature of moral love and hatred.<sup>56</sup> In particular, of Hutcheson's providentialist explanation of moral love and hatred the products of an innate moral sense.<sup>57</sup> I cannot argue the point here, but I think that a similar thing is true of Hume's account of what he calls 'the mixture of benevolence and anger with compassion and malice' (T 2.2.9, SBN 381–9). Here too we find that Hume faces systematic explanatory difficulties in accounting for morally salient other-directed passions. In particular, Hume must explain why our sympathising with the sufferings of strangers often leads us to love them and act benevolently – rather than to hate them, as his double relation theory predicts we should (T 2.2.9.11, SBN 384–9). Hume's solutions here require him to stretch his system of the passions to the limits of its credibility, but, whatever their merits, I think that it is safe enough to assert that Hume intended his arguments on this head to oppose Hutcheson's explanation of compassionate benevolence as the product of an innate 'public sense', and his providentialist denial of genuinely disinterested malice.<sup>58</sup>

What I further maintain, however, is that the difficulties that Hume faces in explaining these morally salient other-directed passions are evidence that his intentions in doing so were secondary ones. More precisely, we find that Hume repeatedly faces difficulties in accounting for benevolent passions within his system of the passions because that system was originally designed to account for the self-interested passions at the origin of civil government. This is not to say, to be clear, that Hume's attempt to account for benevolent concerns within his system of the passions is less important than his attempt to account for the passionate conflicts engendered by riches and property. My claim, rather, is that Hume's chief conceptual devices – sympathy, comparison, and double relations of ideas and impressions – were originally invented to

56 See here Elizabeth S. Radcliffe, "Love and Benevolence in Hutcheson's and Hume's Theories of the Passions" *British Journal for the History of Philosophy*, 12(4) (2004): 631–653.

57 See here also Taylor, 1–31; Merivale, 49–69.

58 Merivale makes similar points, although he sees Hume's chief target as Joseph Butler. See Merivale, 49–50.

the end of explaining the origin of government, and only later put to the purpose of explaining our natural benevolence. This is why – I contend – we find that Hume’s system of the passions, considered as a system, is so well-suited to explaining the origin of avarice and ambition, and so ill-suited to explaining the origins of our benevolent concerns for those around us.

Reading Hume’s philosophy of the passions as having been shaped by primary and secondary intentions helps to explain some of the distractedness of Book Two. Hume’s ad hoc introduction of ‘blended’ passions into his theory of love and hatred, and his inclusion of benevolence and anger amongst the ‘direct’ passions, can both be understood as his making adjustments to a system of the passions that he originally invented for a purpose other than accounting for our natural benevolence. Neither of these moves is incompatible with his general system, but both are necessary to render benevolence explicable. As such, one feels that if Hume had begun designing his system of the passions with the intention accounting for *both* the conflicts engendered by riches and property *and* our natural benevolence, he could have built these adjustments into his system from the outset. Indeed, it is arguable that the main conceptual adjustments that Hume makes in the *Dissertation* serve exactly this purpose.<sup>59</sup>

At the same time, reading Hume’s philosophy of the passions as the product of two distinct intentions allows us to ask whether there is an underlying order to Book Two; such that Hume might be seen to have some further end in mind when explaining both the selfish origin of government, as well as our natural benevolence, within the same system of the passions. I am inclined to believe that Hume did have a larger design for his philosophy of the passions, although I admit that my thoughts here are rather speculative.

Namely, it seems to me that insofar as Hume was concerned to explain the psychology behind why government becomes ‘absolutely necessary’ once a society is affluent, he was concerned to do so without supposing any essential differences between humans and animals. Instead what Hume attempts to show is that all the passions he draws upon to explain violent competition for riches and possessions, including the will, are of the same nature in animals, and arise from the same principles: ‘making a just allowance for our superior knowledge and understanding’ (T 2.1.12.5, SBN 326). Thus each of the three parts of Book Two ends with Hume observing that the phenomena he has just examined are of the same nature in animals, and are ‘excited by the same causes as in human creatures’ (T 2.3.9.32, SBN 448). *Of pride and humility*, ends with the section ‘*Of the pride and humility of animals*’ (T 2.1.12, SBN 324–8); ‘*Of love and hatred*’, with the section ‘*Of the love and hatred of animals*’ (T 2.2.12, SBN 397–8); while ‘*Of the will and direct passions*’ ends with Hume deciding to ‘wave (sic) the examination of the will and direct passions, as they appear in animals; since nothing is more evident, than that they are of

59 See here Merivale, 70–90.



the same nature, and excited by the same causes as in human creatures' (T 2.3.9.32, SBN 448).

Without wanting to say too much about why Hume sought to do this, I think it is reasonable to believe that Bernard Mandeville was a key influence here. For Mandeville himself had held that:

the Fitness of Man for Society, beyond other Animals, is something real; but that it is hardly perceptible in Individuals, before great Numbers of them are joynd together, and artfully manag'd. Secondly, that this real Something, this Sociableness, is a Compound, that consists in a Concurrence of several Things, and not in any one palpable Quality, that Man is endued with, and Brutes are destitute of.<sup>60</sup>

In particular, I think that Hume – following Mandeville – sought to trace the origins of human sociability (and especially the establishment of civil government) to qualities of human nature that we share we with animals. But that – unlike Mandeville – his aim in doing so was not to debase human nature to the level of animals, but to develop a theory of human sociability that an impartial moralist, arguing only from what they can observe, could accept as probably true.<sup>61</sup>

So far as this was part of Hume's design when first developing his system of the passions, we can also see this intention operating behind his attempt to account for our natural benevolence. More precisely, it would seem that Hume's desire to account for benevolent behaviour within the same system of the passions that he had designed to account for the establishment of civil government was also part of an attempt to trace the origins of human sociability to qualities that we share with animals. Towards this point, it is notable that Hutcheson had attributed our benevolence towards those whom we love for their moral qualities, and our good-will towards those we pity, to essential differences between human and animal natures, in the form of innate moral and public senses.<sup>62</sup>

What Hume hoped to gain from this impartial presentation of mankind's sociable qualities is a question unto itself. Still, I think it safe to conclude that the unifying intention behind Hume's philosophy of the passions is to trace the origins of human sociability to qualities that we share we with animals. Hume achieves this by arguing that his explanations of the desire for riches and possessions, as well as benevolence, appeal only to passions and principles that can be found to operate throughout 'the whole sensitive

---

60 Bernard Mandeville, *The Fable of the Bees or Private Vices, Publick Benefits*, 2 vols., ed. F.B. Kaye (Indianapolis: Liberty Fund, 1988), vol 2, 209–10.

61 For more on this point see P.J.E. Kail, "The Sceptical Beast in the Beastly Sceptic: Human Nature in Hume", *Royal Institute of Philosophy Supplement*, 70 (2012), 219–31.

62 Francis Hutcheson, *An Essay on the Nature and Conduct of the Passions and Affections, with Illustrations on the Moral Sense*, ed. Aaron Garrett (Indianapolis: Liberty Fund, (2002), 90.

creation' (T 2.2.12.1, SBN 397); such that '[e]very thing is conducted by springs and principles, which are not peculiar to man, or any one species of animals' (T 2.2.12.1, SBN 397).

But, on second thoughts, how safe is this conclusion? For if such an intention is supposed to underlie the whole of Book Two, how then ought we to understand Hume's account of curiosity, with which he concludes his account of the passions? Hume's treatment of curiosity, or more exactly, of 'the love of knowledge, which displays itself in the sciences' (T 2.3.10.11, SBN 453) is notable for sitting somewhat apart from the rest of Hume's theory of the passions.<sup>63</sup> Whereas parts one and two of Book Two end with an examination of analogous animal passions, part three ends with a waiver of such analogies regarding 'the will and direct passions, *as they appear in animals*' (T 2.3.9.32, SBN 448; my italics) and instead concludes with an account of 'curiosity, or the love of truth' (T 2.3.10, SBN 448–54). Clearly Hume did not think that a love of knowledge was common to man and beast, yet such scientific curiosity is obviously something that renders human beings sociable creatures. Thus it would seem, *prima facie*, that this part of our sociable nature, our love truth, is explained by some quality of human nature that marks a difference in kind between human beings and animal natures. Namely, an innate desire to know. Indeed, this had been the standard position on the nature of curiosity since at least Aristotle, all the way up to Hobbes and beyond.<sup>64</sup>

Nor is this the only interpretive problem posed by Hume's account of curiosity. For it would seem that Hume realised rather late in the game that he ought to examine curiosity within his theory of the passions. In fact he opens his account of this passion with the claim that:

methinks we have been not a little inattentive to run over so many different parts of the human mind, and examine so many passions, without taking once into the consideration that love of truth, which was the first source of all our enquiries. 'Twill therefore be proper, before we leave this subject, to bestow a few reflections upon that passion, and shew its origin in human nature. 'Tis an affection of so peculiar a kind, that t'would have been impossible to have treated of it

---

63 Hume distinguishes between 'the love of knowledge, which displays itself in the sciences' and 'a certain curiosity implanted in human nature, which is a passion derived from a quite different principle' (T 2.3.10.11, SBN 453). I take this second sort of curiosity to be a mode of what Hume calls the '*original* instinct' by which the mind 'tends to unite itself with the good, and to avoid the evil' (T 2.3.9.2, SBN 438). In this case the good desired is the pleasant feeling of stable beliefs, while the evil to be avoided is the painfulness of uncertainty (see here T 2.3.10.12, SBN 453–4). The expression of this desire as a desire to know may be unique to humans (but this is far from clear, animals can be quite inquisitive) but the principle from which it is derived is surely common to both, at least by Hume's lights.

64 For a general overview here see Lorraine Datson and Katharine Park, *Wonders and the Order of Nature, 1150–1750* (New York: Zone Books, 1998).

under any of those heads, which we have examin'd, without danger of obscurity and confusion.

(T 2.3.10.1, SBN 448)

That Hume originally intended to include an account of curiosity within his theory of the passions and then simply forgot to do so is unbelievable. Yet that Hume did not have curiosity in mind when first developing his theory of the passions, and only came to see the importance of accounting for it after having already formulated his division of the passions into the 'direct' and 'indirect' would be entirely in keeping with the nature of Book Two. In fact I am inclined to believe that Hume's skeptical crisis at the end of Book One, which he attempts to resolve by appealing to both curiosity and ambition (see T 1.4.7.12–3, SBN 270–2), brought it home to him that he would need to account for what he calls 'the love of knowledge, as it displays itself in the sciences' within the same system of the passions that he had already developed to account for ambition.

Could it be, then, that Hume's late need to provide an account of curiosity throws out his general project of tracing the origins of human sociability to qualities that we share we with animals? Maybe. But more likely not. Again I cannot argue the point in detail, but I think that if we attend closely to Hume's account of 'the love of knowledge, which displays itself in the sciences' then we find that Hume attempts to show that this passion 'is conducted by springs and principles, which are not peculiar to man, or any one species of animals' (T 2.2.12.1, SBN 397).

Notable here is that Hume's first move is to assert that curious endeavour is not a desire for knowledge 'merely as truth, and that 'tis not the justness of our conclusions, which alone gives the pleasure' of discovering what we believe to be truths (T 2.3.10.2, SBN 448–9). Rather, Hume compares the nature of curiosity to that of hunting, insofar as the 'principal foundation of the pleasure' that we derive from curious endeavour is 'not deriv'd originally from the end, but merely from the action and pursuit' (T 2.3.10.7, SBN 451). At the same time, however, the pleasure of the chase is not enough, Hume thinks, to keep us focused on our game, either when we hunt for sport or for knowledge. Instead, 'these actions must be attended with an idea of utility, in order to their having any effect upon us' (T 2.3.10.8, SBN 451–2). Such ideas of utility, Hume argues, are the product of our sympathizing with those whom we believe will benefit from our pursuits. A sympathy that even if rather 'remote', still serves to 'fix our attention' on the object of pursuit, be it a stag, or the principles of human nature (T 2.3.10.6, SBN 450–1).

Hume's explanation of the desire for knowledge as arising from the 'same principles' (T 2.3.10.9, SBN 452) as the desire to hunt can be read as an attempt to explain this uniquely human passion using principles found to operate throughout 'the whole sensitive creation' (T 2.2.12.1, SBN 397).

Towards this point, it is pertinent that Hume had previously observed, when examining the love and hatred of animals, that:

*sympathy*, or the communication of passions, takes place amongst animals, no less than amongst men ... Every one has observ'd how much more dogs are animated when they hunt in a pack, than when they pursue their game apart; and it is evident that this can proceed from nothing but sympathy. 'Tis also well known to hunters, that this effect follows to a greater degree, and even in too great a degree, when two packs, that are strangers to each other, are join'd together. We might, perhaps, be at a loss to explain this phænomenon, if we had not experience of a similar in ourselves.

(T 2.2.12.6, SBN 398)

On a surface level reading of this passage the similar experience that we have is of hunting for sport with strangers. Yet on deeper reading, one that is justified by the close analogy that Hume draws between hunting and curiosity, we can understand Hume as referring to the disparate factions of the republic of letters, collectively searching after truth, sometimes with too much zeal. So far as this is a good reading of the intentions behind Hume's account of curiosity, we can conclude that Book Two – all of it, not just some of it – is unified by the intention to trace the origins of human sociability to qualities that we share we with animals.

But still we cannot rest easy. For it is at least arguable that because Hume denies that the love of truth is an innate desire to know – and instead holds that this passion arises from a conjunction of the 'same principles' that produce the desire to hunt – Hume obscures his original distinction between passions that are 'indirect' and those that are 'direct'. For what curiosity turns out to be is a 'direct' passion for entertainment, one that is indifferent to its objects, so long as they are entertaining, but which indirectly becomes a love for truth 'itself' through what Hume vaguely describes as 'the natural course of the affections' (T 2.3.10.7, SBN 451): a course which includes our sympathising with the beneficiaries of our discoveries, be they strangers, friends, or enemies.

It is with good reason then that Hume describes curiosity as 'an affection of so peculiar a kind, that t'would have been impossible to have treated of it under any of those heads, which we have examin'd, without danger of obscurity and confusion' (T 2.3.10.1, SBN 448). For the 'love of knowledge, as it displays itself in the sciences' makes no neat fit within a system that begins with the assertion that '[w]hen we take a survey of the passions, there occurs a division of them into *direct* and *indirect*' (T 2.1.1.4, SBN 276). Instead curiosity, properly understood, is both. It is a 'blended' direct-indirect passion that threatens throw the rest of his explanatory system into 'obscurity and confusion'.

As such, what we ultimately find is that while Hume can explain the origin of scientific curiosity using principles that he had already observed to operate in animals, his doing so threatens the coherence of the major conceptual distinction upon which his entire system rests. This, to my mind, is Book Two in a nutshell: an inquiry into the true nature of human sociability that unravels from the pull of its own explanatory ambitions.

#### 4. Conclusion

‘Never literary attempt was more unfortunate than my Treatise of Human Nature’<sup>65</sup>, Hume wrote in 1776, towards the very end of his life. ‘It fell *dead-born from the press*, without reaching such distinction, as even to excite a murmur among the zealots’<sup>66</sup>. The sixty-five year old Hume blamed the still-birth (in his opinion) of his *Treatise* on his youthful ‘indiscretion, in going to press too early’<sup>67</sup>, at the age of twenty-six. Indeed, not long before, Hume had instructed that all further editions of his *Enquiry concerning Human Understanding* and *Enquiry concerning the Principles of Morals* were to come with an ‘advertisement’ to the effect that these works ‘may alone be regarded as containing his philosophical sentiments and principles’<sup>68</sup>. For the *Treatise*, he admitted there, was a ‘juvenile work’, which he had ‘projected before he left college, and which he wrote and published not long after’<sup>69</sup>.

A ‘juvenile work’ is a good description of Book Two. If anything it is more true for Book Two than for Book One or Book Three. Hume’s youthful explanatory incontinence, I think, goes a long way to explaining why Book Two was initially met with such disdain, even by Hume’s champions. Although now, almost seventy years of scholarship later, we have a better sense of the value and importance of Book Two as a part of Hume’s science of human nature. An even greater appreciation, perhaps, than that which the later Hume had for Book Two, given that he never gave his philosophy of the passions the kind of thorough re-working that he afforded the other books of the *Treatise*.

My aim in the above has been to reconcile our present sense of the philosophical importance of Book Two with the fact that it is indeed a juvenile work when read on its own terms. I have argued that Hume’s philosophy of

---

65 David Hume, “My Own Life”, in David Hume, *An Enquiry Concerning Human Understanding*, ed. Stephen Buckle (Cambridge: Cambridge University Press, 2007) (*My Own Life*), 215.

66 Hume, *My Own Life*, 216.

67 Hume, *My Own Life*, 217.

68 See David Hume, *An Enquiry Concerning Human Understanding*, ed. Stephen Buckle (Cambridge: Cambridge University Press, 2007), 2.

69 Hume, *An Enquiry Concerning Human Understanding*, 2.

the passions is the product of three distinct intentions: an attempt to account for the establishment of government, for our natural benevolence, and to trace human sociability back to qualities that are common to the whole of sensitive creation. Accounting for the origins of both government and benevolence was of equal importance to Hume, but I have argued that his accounting for the origin of government was the primary concern of Book Two, and that Hume found himself having to stretch his theory of the passions to accommodate benevolence. Further, I have suggested that Book Two has a distinctive unity of purpose when the first and second of these intentions are subsumed under the third, but that, by the same measure, Hume's desire to place human sociability on a continuum with animal sociability ultimately threatens the coherence of his whole system. We ought not lament this, however. Rather, if we are to understand the distracted and distracting nature of Hume's philosophy of the passions – which is as much a part of its character as its ambitions to systematicity – then perhaps the best way to think of Book Two is as an unstable explanatory compound: a work that fizzes through a series of different explanatory designs, before dissolving into the river of Hume's mature philosophy.

## HUME'S SENTIMENTALISM: NOT NON-COGNITIVISM

**Abstract:** *This paper considers and argues against old and recent readings of Hume according to which his account of moral judgement is non-cognitivist. In previous discussions of this topic, crucial metaethical distinctions—between sentimentalism and non-cognitivism and between psychological and semantic non-cognitivism—are often blurred. The paper aims to remedy this and argues that making the appropriate metaethical distinctions undermines alleged support for non-cognitivist interpretations of Hume. The paper focuses in particular on Hume's so-called 'motivation argument' and argues that it is a poor basis for non-cognitivist interpretations. While there is textual support for attributing to Hume what may be called 'modally weak' motivational internalism, there is no solid textual support for attributing to him either psychological or semantic non-cognitivism. The paper also challenges briefly some further alleged support for non-cognitivist interpretations. It concludes by offering some positive evidence against such interpretations, namely that Hume appears to hold that there are moral beliefs and moral knowledge.*

**Keywords:** *Cognitivism, Hume, internalism, motivation, non-cognitivism*

### 1. Introduction

Non-cognitivist interpretations of Hume's metaethics have for several decades been rather popular both inside and outside of Hume scholarship (see, e.g., Flew 1963; Fogelin 1979; Snare 1991; Bricke 1996). My main aim in this paper is to challenge and undermine such interpretations. I do so by distinguishing clearly between Hume's sentimentalism and different kinds of non-cognitivism. I also argue that the so-called 'motivation argument' is a poor basis for non-cognitivist interpretations. In the light of some recent contributions to Hume scholarship in which non-cognitivist interpretations are severely criticized, it may appear that pursuing the project of this paper is no more worthwhile than flogging a dead horse (see, e.g., Radcliffe 2006; Botros 2006; Cohon 2008; Sturgeon 2008; Sayre-McCord 2008; Lo 2009; Pigden 2009). There are two reasons for why this is only an appearance. First, in scholarly discussions about Hume's metaethics, crucial metaethical distinctions between kinds of non-cognitivism are often blurred. As a consequence, not all ways in which Hume might be read as a non-cognitivist about ethics are properly canvassed, and correlatively, the reasons why Hume

is not to be read as a non-cognitivist remain unclear. Secondly, non-cognitivist interpretations have recently received renewed support (Joyce 2009; Smith 2009; Sobel 2009; Shecaira 2011; Chamberlain 2019).<sup>1</sup> I shall argue that making the appropriate metaethical distinctions undermines this renewed support and shows that non-cognitivist interpretations remain unpersuasive.

I begin in section 2 by clarifying a crucial distinction that is often blurred, between *psychological* and *semantic* non-cognitivism. In section 3, I distinguish these two kinds of non-cognitivism from Hume's *sentimentalism* about ethics. In sections 4–5, these distinctions are put to use in order to undermine non-cognitivist interpretations of Hume and to respond to the aforementioned recent attempts to revive and defend such interpretations. In section 6, I offer some textual evidence that does not merely undermine but count positively against non-cognitivist interpretations.

Before I begin I shall comment on a worry about anachronism. The question whether Hume was a non-cognitivist about ethics may seem absurd. After all, the terms 'cognitivism' and 'non-cognitivism' appeared no earlier than in twentieth century debates, and it is arguable that at least non-cognitivist theories are also late modern inventions. But we can sensibly ask which one, if any, of the modern metaethical views *fits best* with Hume's theory as a whole. Pursuing this question is both historically and philosophically interesting and illuminating, since it promises to enrich our understanding of Hume's theory as well as of modern metaethics.<sup>2</sup>

## 2. Non-Cognitivism: Psychological and Semantic

Non-cognitivism about ethics is often said to be a theory of moral judgement. However, this description is imprecise, due to the ambiguity of the expression 'moral judgement'. First, it can mean the psychological state of holding a view on some moral matter, for example, the view that benevolence is a virtue. According to non-cognitivism about this kind of psychological state, to hold the view that benevolence is a virtue is not (primarily) to have a belief about benevolence, but to hold a non-cognitive attitude to benevolence, for example, to approve of it. Let us call this view *psychological non-cognitivism*. And let us call the view that to hold a view on some moral matter is (primarily) to have a belief about that matter, *psychological cognitivism*. I say that according to psychological non-cognitivism, to hold a moral view is not *primarily* to have a belief, because on some versions of non-cognitivism—

1 It should be noted that while Smith and Shecaira both argue that Hume is best read as a non-cognitivist, Joyce's reading of Hume's metaethics as involving non-cognitivist strands is highly tentative, and Sobel offers his interpretation as a rational reconstruction of what Hume says. Chamberlain attributes to Hume an 'emotivist' theory of moral judgement.

2 Similar remarks about the sensibility and interest of interpreting Hume's metaethics in terms of modern metaethical categories have been made by Cohon (2008: 96); Radcliffe (2006: 358; 2018: 138); Shecaira (2011: 268); and Sturgeon (2008: 513–14).



for example, R. M. Hare's prescriptivism (Hare 1952) and more recent hybrid theories (e.g., Ridge 2007)—to hold a moral view is not only to hold a non-cognitive attitude but also (secondarily) to have a belief. Importantly, the beliefs that are, on these views, partly constitutive of moral judgements have only *non-moral* content. For non-cognitivists, there is no such thing as a belief with *moral* content. This is because there is, on these views, no such thing as moral content of attitudes; what we may think and speak of as 'moral beliefs' are in fact desires or desire-like attitudes whose content is non-moral.<sup>3</sup>

Secondly, 'moral judgement' can mean the linguistic act of uttering a moral sentence, for example, 'Benevolence is a virtue.' According to non-cognitivism about this kind of linguistic act, its meaning is not to be understood truth-conditionally in terms of a proposition it expresses. For, at least on traditional versions of non-cognitivism, such an utterance has no truth-conditions and expresses no (moral) proposition.<sup>4</sup> Instead, the meaning of moral utterances is to be understood in terms of the psychological state (e.g. approval or disapproval) that the uttered sentence (e.g., 'Benevolence is a virtue') conventionally expresses. Let us call this view *semantic non-cognitivism*. And let us call the view that the meaning of a moral utterance is to be understood in terms of its truth-conditional content, that is, the proposition it expresses, *semantic cognitivism*.

For the purposes of this paper, it is important to notice that psychological non-cognitivism does not entail semantic non-cognitivism.<sup>5</sup> To illustrate, one might hold that to think that benevolence is a virtue is to approve of benevolence, and that to utter the sentence 'Benevolence is a virtue' is not to *express* approval of benevolence but to *report* that one approves of benevolence. On this view, a speaker's utterance of 'Benevolence is a virtue' expresses a proposition and has a truth-condition, namely that the speaker approves of benevolence. This would be to combine a version of psychological non-cognitivism with a version of semantic cognitivism. Such a combination of views is perfectly coherent. Whether it is also an attractive metaethical view in its own right is a further issue that will not detain us here. The main points here—which are crucial for the debate on how to interpret Hume's metaethics—are that psychological and semantic non-cognitivism are distinct positions and consequently that arguments that support the one need not support the other.

3 I take it that this is what Michael Smith has in mind when he says that according to some non-cognitivists, moral beliefs are 'constituted by desires' (Smith 2009: 109).

4 I say that on traditional versions of non-cognitivism, utterances of moral sentences express no *moral* propositions, since on some traditional versions of non-cognitivism, for example, Hare's, moral utterances have a secondary descriptive meaning in that they express descriptive (non-moral propositions); see Hare 1952. On some recent, non-traditional versions of non-cognitivism, moral utterances have *deflationary* truth-conditions.

5 Neither does semantic non-cognitivism entail psychological non-cognitivism. But the combination of semantic non-cognitivism and psychological cognitivism is more exotic than the combination of psychological non-cognitivism and semantic cognitivism.

Yet the distinction is often blurred in discussions of Hume's metaethics, by advocates as well as by critics of non-cognitivist interpretations. For example, Fábio Shecaira, who defends a non-cognitivist interpretation, takes non-cognitivism to be a 'semantic theory of moral statements' (Shecaira 2011: 268, 272), and he goes on to say that according to non-cognitivism, 'moral judgements are not beliefs (they are not truth-evaluable and, therefore, are not generated by reason), but they should instead be equated with emotions, valuations, avowals of prescriptions, or some other kind of practical, non-representational attitude' (Shecaira 2011: 275). Rachel Cohon, who is critical of non-cognitivist interpretations, takes non-cognitivism to be the view that '[m]oral judgements are not cognitive states or representations, but mere feelings or expressions of feeling; they do not assert propositions or represent states of affairs, and can be neither true nor false' (Cohon 2008: 11).

Both of these descriptions blur the distinction between psychological and semantic non-cognitivism. As we have seen, psychological non-cognitivists maintain that moral judgements are not (primarily) beliefs or cognitive or representational states; they are rather to be equated with certain kinds of emotions or feelings. But psychological non-cognitivists do not have to agree that moral utterances are not truth-evaluable and do not assert propositions; they can, maintain, for example, that moral utterances report the attitudes (emotions or feelings) that constitute moral judgement. In other words, psychological non-cognitivists can accept semantic cognitivism. This point is important since much of the alleged evidence for non-cognitivist interpretations of Hume supports only *psychological* non-cognitivist interpretations. We have seen that such evidence need not support *semantic* non-cognitivist interpretations.<sup>6</sup>

### 3. Hume's Sentimentalism

Next, we need to consider Hume's *sentimentalism* and how it relates to the two kinds of non-cognitivism we have distinguished. Hume's foremost metaethical question is 'whether it be possible, from reason alone, to distinguish betwixt moral good and evil' (T 3.1.1.4; SBN 457). That is to say, Hume's foremost metaethical focus is on the *epistemological* question of how we distinguish virtue from vice, good from evil, and right from wrong. The view he is primarily concerned to refute is the *rationalist* view that we can do so by means of reason alone. He endorses the *sentimentalist* view that sentiments of approbation and disapprobation are what ultimately enable us to make moral distinctions.

---

6 James Chamberlain has recently argued that Hume holds an 'emotivist' theory of moral judgement (Chamberlain 2019). Like many others, however, Chamberlain fails to distinguish semantic from psychological accounts. It is not clear whether he takes emotivism to be a semantic or a psychological account of moral judgement, or both (see, e.g., pp. 1059–60, 1070–71).

Hume famously defines virtue as 'whatever mental action or quality gives to a spectator the pleasing sentiment of approbation, and vice the contrary' (EPM App. 1.10; SBN 289). He equally famously holds that reason is concerned with two things only, namely relations of ideas and matters of fact concerning cause and effect (EHU 4.1–4.2; SBN 25–26; see also T 2.3.3; SBN 413–18). Reason alone does not suffice to distinguish virtue from vice, because reason alone cannot *originally* tell us which kinds of mental actions or qualities give rise to sentiments of approbation or disapprobation. In order to make moral distinctions, we need to have experiences of instances of the relevant mental actions and qualities and of the sentiments they give rise to (Cohon and Owen 1997; Kail 2007: 192–3). That is the sense in which '[m]oral distinctions [are] not derived from reason', as the title of section 3.1.1 of the *Treatise* states. Once we have such experiences, reason can inform us of their causes and since 'like causes always produce like effects' (T 1.3.15.8; SBN 174), reason can help us realize that similar mental actions and qualities are likely to give rise to similar sentiments, and are therefore likely to be virtues or vices, depending on the kind of sentiment they give rise to.

For example, reason alone cannot inform us that benevolence is a virtue. We need to have experience of a real or imaginative benevolent character and the pleasing sentiment of approbation that arises from such an experience. Once we have had such experiences and sentiments we can reason from them to the conclusion that similar benevolent characters are, or would be, virtuous too. More generally, reason alone cannot tell us which kinds of activities are pleasant or agreeable and which are not. To give a non-moral example, experience may tell us that drinking cold beer on a hot summer day is pleasant, and reason can help us infer from such experiences that drinking similar liquids in similar circumstances is likely to be pleasant too. Now, if the inability of reason alone to inform us about what character traits are virtues or vices supports non-cognitivism about moral judgements, then the inability of reason alone to tell us which kinds of activities are pleasant and which are unpleasant should, in the same manner, support non-cognitivism about judgements concerning the pleasant and the unpleasant.<sup>7</sup> But the latter view is highly unattractive and to my knowledge, no one has defended it. By the same token, the mere claim that reason alone does not suffice to make moral distinctions—distinctions between virtue and vice—gives no support to non-cognitivism about ethics.

Since Hume's sentimentalism is foremost an *epistemological* view of moral judgement it has no direct implications concerning *semantic* non-cognitivism and cognitivism. However, it might be tempting to think that Hume's sentimentalism fits better with *psychological* non-cognitivism than with psychological cognitivism. In order to resist that temptation, let us suppose that one holds the psychological cognitivist view that to judge that

---

7 Pigden (2009: 96) makes a similar point.

benevolence is a virtue is to believe that benevolence gives rise to sentiments of approbation in spectators. That view is perfectly compatible with the epistemological view that the belief that benevolence gives rise to approbation in spectators cannot be based on reason alone, but must ultimately be based on impressions of the relevant kinds of sentiments or passions; just as my belief that I am now hungry, or that drinking cold beer on a warm day is pleasant, must be based on the relevant kinds of sense impressions.<sup>8</sup> These beliefs are all ultimately non-inferential.

The fact that Hume's sentimentalism is perfectly compatible with psychological cognitivism does not imply that it is *incompatible* with psychological non-cognitivism. It has been noted before that Hume's sentimentalism is compatible with a number of different metaethical views (e.g., Mackie 1980: 73–5). We shall see in section 5 below that some possible readings of Hume as a non-cognitivist rely on failures to distinguish appropriately between sentimentalism and non-cognitivism, but the most notable non-cognitivist interpretations of Hume are not based simply on his sentimentalism but on more specific claims and arguments he makes. One such argument that has been thought to provide strong support to non-cognitivist interpretations is the motivation argument, to which we turn next.<sup>9</sup>

#### 4. The Motivation Argument: Not Non-Cognitivist

Hume famously holds that 'morals [...] have an influence on [...] actions and affections' (T 3.1.1.6; SBN 457). In contrast, 'reason [is] cool and disengaged [and] is no motive to action' (EPM App. 1.21; SBN 294). A canonical statement of what has become known as Hume's 'motivation argument' is this (at T 3.1.1.6; SBN 457):

- (M) 'Morals excite passions, and produce or prevent actions.'
- (R) 'Reason of itself is utterly impotent in this particular.'
- (C) 'The rules of morality, therefore, are not conclusions of our reason.'

As many commentators have pointed out, the argument is far from obviously valid. In order to render it valid, we might try the following reformulation:

- 
- 8 Elizabeth Radcliffe distinguishes between cognitivism and non-cognitivism concerning the 'process' of judging morally and its 'product', that is, moral judgement (Radcliffe 2006: 361, 363; see also Radcliffe 2018: 142–43). The process of judging morally, as she describes it, is broadly speaking an epistemic process, so her distinction between non-cognitivism and cognitivism about this process corresponds to my distinction between sentimentalism and rationalism. With respect to the 'product' of this process—moral judgement—Radcliffe makes no distinction between the psychological state and the linguistic act of making a moral judgement, so she does not distinguish as I do, between psychological and semantic cognitivism and non-cognitivism.
  - 9 Other traditional grounds for non-cognitivist interpretations of Hume, such as the is-ought passage, are amply criticised in Sturgeon 2008 and Pigden 2010. See also Cohon 2008.

- (M\*) Moral judgements excite passions and produce or prevent actions.
- (R\*) Conclusions of our reason cannot of themselves excite passions or produce or prevent actions.
- (C\*) Moral judgements, therefore, are not conclusions of our reason.<sup>10</sup>

The inference from (M\*) and (R\*) to (C\*) is still not obviously valid. For it might be that moral judgements are conclusions of our reason and that we are so disposed as to be motivated to act in accordance with our moral judgements. Without this disposition, however, moral judgements would not motivate, since conclusions of reason in and of themselves cannot motivate. In that way, (M\*) and (R\*) could both be true and (C\*) false.

In order to attribute to Hume a valid argument, it is a standard suggestion that we read (M) and (M\*) as expressing a version of motivational *internalism* (Radcliffe 2006: 356; Shecaira 2011: 269). Motivational internalism is ordinarily understood as the claim that it is necessary that if an agent makes a sincere moral judgement, she is to some extent motivated to act in accordance with that judgement.<sup>11</sup>

However, not everyone understands internalism in this way. In his recent non-cognitivist interpretation of Hume, Fabio Shecaira takes internalism to be the view that moral judgements are *intrinsically* motivating. That is to say, moral judgements are capable of motivating on their own, 'without the contribution of an independent state of mind' (Shecaira 2011: 269, 277; Radcliffe 2018: 115–16), such as a disposition or standing desire to act in accordance with one's moral judgements. We can call this version of internalism 'modally weak', since it allows that moral judgements do not motivate necessarily. For even if moral judgements can motivate intrinsically, this motivation can be blocked by independent states of mind.

In the context of interpreting Hume, the modally weak version of internalism is significant in a way that deserves to be emphasized. For it turns out that in order to render the motivation argument valid, we need not interpret premise (M\*) as the claim that moral judgements motivate *necessarily*. Since premise (R\*) says that conclusions of reason of themselves *cannot* motivate, it suffices that premise (M\*) says that moral judgements of

10 This formulation of the motivation argument is in line with Shecaira's interpretation (2011: 269–70). It is not an uncontroversial interpretation, however. Rachel Cohon has argued that the motivation argument concerns moral properties rather than moral judgements (Cohon 1997: 259–61). I find her reading doubtful on this score, but shall not consider it further here, since I want to grant as much as possible to defenders of non-cognitivist interpretations.

11 'Internalism' is used as a label for a regrettably large number of different positions in philosophy. In what follows, I shall use it only as a label for a distinctive view about the connection between moral judgement and motivation to act. For expository reasons, I shall mostly drop the qualifier 'motivational'.

themselves *can* motivate. In other words, in order validly to derive Hume's anti-rationalist conclusion that moral judgements are not conclusions of reason, we need not attribute to Hume internalism as ordinarily understood, but merely the view that moral judgements can motivate on their own (modally weak internalism) and that conclusions of reason cannot motivate on their own.<sup>12</sup>

Shecaira's argument for interpreting Hume as a non-cognitivist is that internalism fits much better with non-cognitivism than with cognitivism (Shecaira 2011: 275). Whether Hume is an internalist in any sense is contested.<sup>13</sup> I believe that modally weak internalism can be plausibly attributed to Hume. But we shall see that even so, the motivation argument still does not support non-cognitivist interpretations of Hume, whether psychological or semantic.

#### 4.1. *Internalism and Semantic Non-Cognitivism*

According to Shecaira, '[s]ome semantic theories about moral statements are incompatible with internalism' (Shecaira 2011: 272). That claim is mistaken. The supposed intrinsic connection between moral judgement and motivation to act obtains between the *psychological state* of making a moral judgement and being motivated to act. It would not be plausible to maintain that the connection holds only between the *linguistic act* of uttering a moral sentence and being motivated to act, for in many contexts we do not utter any moral sentences but are still motivated to act in accordance with what we judge to be morally virtuous or vicious, good or evil, or right or wrong. It would be highly implausible to maintain that in order to secure a connection between the judgement that, for example, benevolence is a virtue, and motivation to praise or emulate benevolent characters, we would have to utter a sentence like 'Benevolence is a virtue'. If anything, that would seem to so show that the supposed connection between the psychological state of judging that benevolence is a virtue and being motivated to act is *not* intrinsic.

Contrary to Shecaira's claim, then, no semantic theories about moral statements are incompatible with internalism, since internalism is a theory about the psychology of moral judgement, not the semantics of moral judgement. But perhaps Shecaira could concede that internalism is not

---

12 Elizabeth Radcliffe interprets Hume as an internalist, but it is not clear whether she means to attribute to Hume the standard version of internalism or merely the modally weak one. Some of her formulations of internalism suggest the latter, while some others suggest the latter. See Radcliffe 2006: 353, 355; 2018: Ch. 5.

13 For example, Charlotte Brown has argued that Hume is inconsistent in that he accepts internalism in the motivation argument but rejects it elsewhere (Brown 1988), while Elizabeth Radcliffe argues that Hume is a consistent internalist (Radcliffe 2006, 2018), as does Shecaira (2011: 276–8). According to Geoffrey Sayre-McCord (2008), there is evidence that Hume was not an internalist and he argues that none of Hume's arguments rely on internalism. Simon Blackburn has recently defended an externalist reading of Hume (Blackburn 2015).

incompatible with semantic cognitivism, and revert to his weaker claim that internalism *fits better* with semantic non-cognitivism. It is not clear, however, exactly what it would be for a semantic theory of the meaning of moral utterances to fit well or badly with a psychological theory of what it is to hold a moral view. There is more plausibility in the claim that internalism fits better with psychological non-cognitivism than with psychological cognitivism, at least given some currently popular views about which kinds of mental states are motivationally efficacious and which are not. As we shall see in the next section, however, Hume did not endorse these currently popular views.

#### 4.2. *Internalism and Psychological Non-Cognitivism*

It might be that reading Hume as an internalist supports attributing to Hume psychological non-cognitivism. Psychological non-cognitivism offers a very simple and straightforward explanation of the supposed intrinsic connection between moral judgement and motivation to act, since it holds that to judge that someone is virtuous or vicious, or that something is good or evil, or right or wrong, simply *is* (or is primarily) to hold a non-cognitive and motivationally efficacious attitude to whoever or whatever it is that is judged virtuous or vicious, good or evil, or right or wrong.

However, this explanation of the supposed connection supports psychological non-cognitivism over psychological cognitivism only on the assumption that no cognitive state, that is, no belief, is motivationally efficacious. It is a popular view in modern metaethics that beliefs are motivationally inert, whereas desires are motivationally efficacious: In order to be motivated to act, an agent must have a desire and a relevant means-end belief. This view often goes under the name—misleadingly, as we shall see—the ‘Humean Theory of Motivation’ and is often based on a functional account of belief and desire, according to which beliefs and desires have different, in fact opposite, directions of fit. The idea is that beliefs are states that aim to fit the world, while desires are states that aim to make the world fit them (Smith 1994: Ch. 3). It is easy to see how and why many who accept internalism and the Humean Theory of Motivation are also attracted to psychological non-cognitivism about moral judgement.

Crucially, however, Hume himself did not accept the Humean Theory of Motivation. In particular, he did not accept the claim that no belief is motivationally efficacious. What he did accept is the much more restricted claim that no belief based on reason alone is motivationally efficacious. But that claim is far too restricted to justify attributing psychological non-cognitivism to Hume.

A potential mistake is to take Hume's famous claim about the motivational inertia of conclusions based on reason alone (T 2.3.3.3–4; SBN 414–15) to concern belief in general. In fact, Shecaira makes this mistake in that he interprets the second premise of the motivation argument (R\*

above) as the thesis that ‘no belief can (alone) motivate to action’ (Shecaira 2011: 280).<sup>14</sup> A possible explanation of why the mistake may be made is the interpretive assumption that only reason generates beliefs, which is clearly a mistaken interpretive assumption. For example, my current beliefs that I am now hungry and that drinking cold beer on a hot summer day is pleasant are both based on sense impressions and experience, and are not (purely) conclusions of reason. In general, since beliefs for Hume are lively ideas, and since ideas are copies of impressions, there is clearly nothing that rules out the existence of beliefs that are copied from impressions and sentiments, and not generated by the motivationally inert faculty of reason.

As an illustration of Hume’s rejection of the view that no belief alone can motivate action, consider Hume’s account of how impressions of pleasure and pain give rise to ideas of pleasure and pain, which produce motivation to act:

An impression first strikes upon the senses, and makes us perceive heat or cold, thirst or hunger, pleasure or pain of some kind or other. Of this impression there is a copy taken by the mind, which remains after the impression ceases; and this we call an idea. This idea of pleasure or pain, when it returns upon the soul, produces the new impressions of desire and aversion, hope and fear, which may properly be called impressions of reflexion, because derived from it. (T 1.1.2.1; SBN 7–8)

When an idea ‘returns upon the soul’ it may be or become so lively as to amount to a belief. The belief that some behaviour or some object would generate pleasure or pain can thus produce the motivationally efficacious states of desire and aversion. As Hume also says,

we find by experience, that the *ideas* of those [pleasant or unpleasant] objects, *which we believe either are or will be existent*, produce in a lesser degree *the same effect* with those impressions, which are immediately present to the senses and perception. The effect, then, of belief is to raise up a simple idea to an equality with our impressions, and bestow on it a like influence on the passions. (T 1.3.10.3; SBN 119, emphases added.)

It is worth stressing that Hume claims that it is ‘the *idea* of pleasure or pain, when it returns upon the soul, [that] *produces* the new impressions of desire and aversion’ and that ‘*ideas* of [pleasant or unpleasant] objects, which we believe either are or will be existent, *produce* in a lesser degree the same effect with [...] impressions.’ Hence Hume’s view appears not to be that beliefs

---

14 Elizabeth Radcliffe does not make this mistake but argues at length that Hume’s view is that no representational state, such as a belief, can motivate on its own and that the so-called Humean theory of motivation is fully compatible with Hume’s own view. (See Radcliffe 2018, esp. Ch. 2.) While I have some doubts about Radcliffe’s claims on this score, her reading is not in conflict with my main contention that there is no solid evidence for non-cognitivist interpretations of Hume (Radcliffe 2018: Ch. 5).



about what is or will be pleasant or painful are motivating only because of *independent* desires to pursue pleasure and avoid pain, at least that is not how he states his view.<sup>15</sup> His view appears rather to be that beliefs about what is or will be pleasant or painful are motivating because they produce desires to act in the relevant ways.<sup>16</sup>

Since Hume holds that observing and possessing virtue are normally pleasant, and correlatively that observing and possessing vice are normally unpleasant (T 3.3.6.6; SBN 620–1; EPM, ch. 9), Hume could argue that moral beliefs (lively ideas of virtue and vice) motivate in a similar way, that is, by producing desire and aversion, approbation and disapprobation. To illustrate, a spectator's belief that St Francis' benevolence is a virtue is a lively idea that is a copy of an agreeable impression of approbation that the spectator feels when contemplating St Francis' benevolence from the common point of view. When the idea that St Francis's benevolence is a virtue 'returns upon the soul' and acquires the vivacity of belief, it produces in the spectator—by virtue of being a representation of the agreeableness felt when contemplating St Francis' benevolence—various motivationally efficacious attitudes, such as love of St Francis, and a desire to emulate his benevolent character and behaviour. Other kinds of beliefs about objects and their causes and effects—

---

15 Rachel Cohon calls that view the 'background impulse model' (2008: 38–40) and she criticises it effectively, both for being in itself less plausible than the view that beliefs about pleasure and pain give rise to desires without the help of backgrounding general desires and for not fitting very well with Hume's apparent claims that it is the beliefs that produce the desires (2008: 45–9). Shecaira complains that Hume's claims about the motivational efficacy of belief in *Treatise* book I are 'remote' in relation to the discussion of morals in book III, and that on the other side of the argument is Hume's 'clear statement that reason [...] cannot alone motivate to action' (2011: 280). In response to the first complaint, it can be argued that while some of Hume's claims about the causal efficacy of belief may be remote from book III in terms of page numbers, the fact that they are made at a very early stage in book I indicates that they are fundamentally important in his psychology of action. Moreover, Cohon has shown that Hume makes claims even in book III that suggest that he takes some beliefs to be motivationally efficacious (Cohon 2008: 18). For example, Hume says that '[a] person may be affected with passion, by supposing a pain or pleasure to lie in an object' (T 3.1.1.12; SBN 459). In response to Shecaira's second complaint it suffices to repeat a point in the main text above, namely that it is a mistake to take Hume's claims about the motivational inertia of conclusion based on reason to concern beliefs in general.

16 Shecaira is thus mistaken to attribute to Hume the view that '[k]nowledge of X's potential to cause happiness will only drive one to pursue X if one has an independent desire to pursue whatever course of action has the potential to cause happiness' (Shecaira 2011: 273). In fact, several authors have noted recently that Hume holds that beliefs about prospects of pleasure and pain can motivate (see, e.g., Cohon 2008: 17–18; Kail 2007: Ch. 8; Sturgeon 2008: 522–3; Pigden 2009: 101). Shecaira considers the point as presented by Sturgeon and he challenges Sturgeon's claim that Hume takes impressions to differ from ideas only in being more forceful and vivacious. According to Shecaira, impressions and ideas also differ in that only the latter are representational and hence capable of truth and falsity (Shecaira 2011: 279). That is correct as far as it goes, but it does nothing to show that Hume holds that no belief is motivationally efficacious

beliefs whose content does not concern prospects for pain and pleasure, and beliefs that do not represent agreeableness or disagreeableness—do not have this kind of motivational force.

This reading settles the question whether any form of motivational internalism can be attributed to Hume. It makes it plausible to attribute to Hume modally weak internalism, according to which moral judgements can motivate intrinsically. For the view is not that moral beliefs motivate by virtue of an *independent* mental state, such as a standing desire or disposition to act in accordance with one's moral judgements; the view is that moral beliefs, produce motivation to act because of their special content: they represent the agreeable or disagreeable sentiments felt when contemplating characters from the common point of view. These representations get associated with ideas of the relevant characters, and so the characters are represented as agreeable (virtuous) or disagreeable (vicious).

To recap, Hume's view that some beliefs can excite passions and produce or prevent actions suffices to show that his motivation argument against moral rationalism is perfectly compatible with reading him as a psychological (and semantic) cognitivist. In other words, this point suffices to undermine the alleged support the motivation argument gives to non-cognitivist readings of Hume. Attributing to Hume modally weak internalism is textually justified, but it does not support non-cognitivist interpretations. In section 6, we shall consider some textual evidence that does not merely undermine non-cognitivist readings but count positively against them. But before we come to that we shall consider some other passages that may be thought to support non-cognitivist interpretations. We shall see that as long as we keep in clear view the distinction between non-cognitivism and sentimentalism, as explained in section 3 above, these passages do not lend support to non-cognitivist interpretations.

## 5. Some Further Alleged Support for Non-Cognitivist Readings

One might be tempted to think that, quite regardless of the best understanding of the motivation argument, Hume's view that we distinguish between virtue and vice by means of our impression rather than merely by means of our ideas, gives some support to non-cognitivist interpretations (T. 3.1.1.3–4; SBN 456–7). But Hume's view that we distinguish vice from virtue by means of impressions support only the sentimentalist view that moral judgements must originally be based in part on sentiments or feelings. The temptation to think that it (also) supports non-cognitivism rests on a failure to distinguish between sentimentalism and non-cognitivism. We saw in section 3 above that there is no incompatibility at all between Hume's sentimentalism and cognitivism.

A similar mistake might lead one to think that the following passage from *Treatise*, book 3, supports non-cognitivism:

To have the sense of virtue, is nothing but to *feel* a satisfaction of a particular kind from the contemplation of a character. The very *feeling* constitutes our praise or admiration. ... We do not *infer* a character to be virtuous because it pleases (T 3.1.2.3; SBN 471, last emphasis added).

These claims may be thought to be in tension with cognitivist interpretations, but in fact they are not. First, to have the *sense* of virtue from the contemplation of a character is not yet to judge that the character is virtuous; the feeling that constitutes the sense of virtue may be, and often is, the basis of such a judgement (Sayre-McCord 2008: 303). Secondly, to praise or admire a character is to have feelings of peculiar kinds that can—and often do—serve as bases of judgements to the effect that the character in question is praiseworthy or admirable. Thirdly, beliefs about the virtue or vice of a character that are based on feelings that result from contemplation of that character are non-inferential beliefs; that is why we do not infer a character to be virtuous because it pleases.

There are also some passages in the second *Enquiry* that might be thought to support non-cognitivist interpretations. Consider Hume's claim that 'when [a speaker] bestows on any man the epithets of *vicious* or *odious* or *depraved*, he ... expresses sentiments, in which, he expects, all his audience are to concur with him' (EPM 9.6; SBN 272).<sup>17</sup> It is plausible, however, that all Hume means to say here is that when we make moral judgements, for example, when we judge that a person's character is vicious, odious, or depraved, we take up the common point of view, and so expect others to concur with our judgement. By contrast, when we express mere dislike of another person, or when we deem him a personal enemy, rival, adversary, or antagonist, we 'speak the language of self-love' (EPM 9.6; SBN 272), and so do not expect others to concur with our judgements. This has no special implications for the truth or falsity of cognitivism, whether psychological or semantic.

One might also suggest that there is support for non-cognitivist interpretations in Appendix 1 of the second *Enquiry*, in which Hume argues that before we can fix on a moral verdict, all the circumstances of the case at hand need to be laid out (Sobel 2009: 65). When they have been, 'the understanding has no further room to operate' and then '[n]othing remains but to feel, on our part, some sentiment of blame or approbation; whence we pronounce the action criminal or virtuous' (EPM App. 1.11–12; SBN 289–91). But this seems to be a description of the process of making a moral judgement, and as such it is fully compatible with cognitivism (psychological and semantic). To think otherwise is again to fail to distinguish appropriately between Hume's sentimentalism and non-cognitivism.

---

17 Richard Joyce (2009: 48) and Howard Sobel (2009: 64) both refer to this passage in their respective tentative readings of Hume's view as involving non-cognitivist strands.

## 6. Evidence against Non-Cognitivist Interpretations

I shall offer two main bits of evidence against non-cognitivist interpretations and I shall begin with a piece of circumstantial evidence. Hume says in the opening chapter of the second *Enquiry* that one question he is interested in pursuing is ‘whether we attain the knowledge of [moral distinctions] by a chain of argument and induction, or by an immediate feeling and finer internal sense’ (EPM 1.3; SBN 170). This formulation clearly indicates that Hume assumes that there is such a thing as moral knowledge.<sup>18</sup> We may safely assume that when Hume intimates that there is moral knowledge, he uses the term ‘knowledge’ in colloquial sense and not in the strict sense stipulated in the *Treatise*, in which knowledge requires certainty (T 1.3.11.2; SBN 124). This is not the only place at which uses the term ‘knowledge’ in a non-strict, colloquial sense. For example, in the first *Enquiry* Hume queries ‘how we arrive at the knowledge of cause and effect’ (EHU 4.5; SBN 27), and in the second *Enquiry*, Hume concludes the section on justice by claiming that his theory provides knowledge of the relevance of public utility to justice and other social virtues (EPM 3.48; SBN 203–4). Here Hume uses ‘knowledge’ in the colloquial and not the strict sense.

The assumption that there is such a thing as moral knowledge sits uneasily with the interpretation that moral judgements are to be equated with passions, since on Hume’s view, passions are non-representational ‘original existences’, which, for that reason, cannot be true or false (T 2.3.3.5; SBN 415).<sup>19</sup> According to Hume, ‘truth [...] consists in an agreement [...] either to the *real* relations of ideas, or to *real* existence and matter of fact [and] ’tis evident our passions [...] are not susceptible to such agreement’ (T 3.1.1.9; SBN 458, Hume’s emphases). Since knowledge implies truth, this suggests strongly that, for Hume, knowledge that some person is virtuous or that some character trait is a virtue is not merely a matter of having a passion towards that person or character trait; it also involves having a lively idea of that person as virtuous, or of that character trait as a virtue. As we saw in the preceding section, this idea can be traced back to an impression of approbation of that person or character trait, and we can say that passions are in that sense grounds of our moral knowledge. But for the reasons just given, moral knowledge also requires or presupposes moral ideas.

The second piece of evidence I will offer is less circumstantial. I said in section 2 above that according to psychological non-cognitivism, to hold a moral judgement is not primarily to hold a belief. If Hume’s view of moral judgement were more in line with psychological non-cognitivism than with psychological cognitivism we would expect him to hold that there are

18 Hume also presupposes that there is such a thing as *political* knowledge (EPM 3.34; SBN 196–7).

19 For an illuminating discussion of Hume’s claim that passions are ‘original existences’, see Cohon and Owen (1997).

no moral beliefs.<sup>20</sup> But I have already intimated—in the present and the preceding sections—that Hume held that there are moral beliefs. One might object that there is no textual support for this reading.

However, on least at one occasion, Hume speaks explicitly about ‘moral ideas’ (‘Of the Immortality of the Soul’, EPML 595) and on several occasions of our ‘ideas of virtue’ and of our ideas of ‘vice and moral deformity’ (T 3.2.2.23; SBN 498, T 3.2.8.7; SBN 545; EPM Dial. 39; SBN 337). Since beliefs for Hume are lively ideas and since there is no textual reason to suppose that Hume held that ideas of virtue and moral ideas cannot be lively, there is no reason not to attribute to Hume the view that there are moral beliefs. Defenders of non-cognitivist interpretations might suggest that Hume speaks loosely in the relevant passages, and that by ‘ideas of virtue’ he rather means *impressions of virtue* or *moral sentiments* (Smith 2009: 112). But such a reading seems strangely contrived and lacks textual support: Why deny that impressions of virtue and vice can give rise to ideas of virtue and vice, where the latter are copies of the former, just as sensory impressions of, for example, colour and heat and cold can give rise to ideas of colour and heat and cold, where the latter are copies of the former? This question has no plausible and textually supported answer.<sup>21</sup>

It is important to note, once again, that none of this compromises Hume’s sentimentalism. For recall that Hume’s sentimentalism maintains that moral judgements cannot be based on reason alone, but must in part be based on sentiments. Reason alone is not capable of making moral distinctions. Once we are aware, by virtue of sentiment, that some character trait gives rise to approbation or disapprobation in a spectator, we can use reason to conclude that similar character traits will give rise to similar sentiments of approbation or disapprobation in a similar spectator. But the original insight that some character trait is a virtue or a vice is due to sentiment and is not a conclusion of reason. In that sense, morality is ‘more properly felt than judg’d of’ (T 3.1.2.1; SBN 470). And in that sense it is not possible ‘*from reason alone*, to distinguish betwixt moral good and evil’ (T 3.1.1.4; SBN 457, emphasis added).

## 7. Conclusion

In this paper we have distinguished between cognitivism and non-cognitivism concerning the psychological state of making a moral judgement, and between cognitivism and non-cognitivism concerning the linguistic act of uttering a moral judgement. We have found that neither kind of non-cognitivism is attributable to Hume. In particular, regardless of whether Hume is a motivational internalist who holds that moral judgements

20 Chamberlain 2019: pp. 1062–64.

21 For another defence of the view that Hume’s view recognizes moral beliefs, see Radcliffe 2018: 138–43.

motivate intrinsically, non-cognitivist interpretations, whether psychological or semantic, remain implausible. Hume is plausibly read as a sentimentalist, but not as a non-cognitivist, about ethics.<sup>22</sup>

## References

- Blackburn, Simon 2015. 'Hume, Morality, and Skepticism', in P. Russell. ed., *The Oxford Handbook of Hume*. Oxford: Oxford University Press.
- Botros, Sophie 2006. *Hume, Reason, and Morality*. London: Routledge.
- Bricke, John 1996. *Mind and Morality*. Oxford: Clarendon Press.
- Brown, Charlotte 1988. 'Is Hume an Internalist?' *Journal of the History of Philosophy* 26: 69–87.
- Chamberlain, J. 2019. 'Hume's Emotivist Theory of Moral Judgements', *European Journal of Philosophy* 28: 1058–72.
- Cohon, Rachel 1997. 'Is Hume a Non-Cognitivist in the Motivation Argument?' *Philosophical Studies* 85: 251–66.
- , 2008. *Hume's Morality: Feeling and Fabrication*. Oxford: Oxford University Press.
- Cohon, Rachel and David Owen 1997. 'Hume on Representation, Reason, and Motivation', *Manuscripto* 20: 47–76.
- Flew, Anthony 1963. 'On the Interpretation of Hume', *Philosophy* 38: 178–82.
- Fogelin, Robert 1979. 'Review of Jonathan Harrison, *Hume's Moral Epistemology*', *Noûs* 13: 523–5.
- Hare, R. M. 1952. *The Language of Morals*. Oxford: Clarendon Press.
- Hume, David 1975 [1748, 1751]. *Enquiries concerning Human Understanding and concerning the Principles of Morals*, 3rd edn., eds. L. A. Selby-Bigge & P. H. Nidditch. Oxford: Clarendon Press.
- , 1978 [1739–40]. *A Treatise of Human Nature*, 2nd edn., eds. L. A. Selby-Bigge & P. H. Nidditch. Oxford: Clarendon Press.
- , 1998 [1751]. *An Enquiry concerning the Principles of Morals*, ed. T. L. Beauchamp. Oxford: Oxford University Press.
- , 1999 [1748]. *An Enquiry concerning Human Understanding*, ed. T. L. Beauchamp. Oxford: Clarendon Press.
- , 2000 [1739–40]. *A Treatise of Human Nature*, ed. D. F. Norton & M. J. Norton. Oxford: Clarendon Press.

---

22 This paper draws on a paper written in Swedish ('Humes sentimentalism—inte non-cognitivism') that appeared in *Filosofisk tidskrift* #3/2018. A research grant from the Swedish Research Council (project id 2019–02828) is gratefully acknowledged.

- , 1994. *Essays, Moral, Political, and Literary*, ed. E. F. Miller. Indianapolis: Liberty Fund.
- Joyce, Richard 2009. 'Expressivism, Motivation Internalism, and Hume', in Pigden, ed., 2009, 30–56
- Kail, Peter J. E. 2007. *Projection and Realism in Hume's Philosophy*. Oxford: Oxford University Press.
- Lo, Norva Y. S. 2009. 'Is Hume Inconsistent?—Motivation and Morals', in Pigden, ed., 2009, 57–79.
- Mackie, John Leslie 1980. *Hume's Moral Theory*. London: Routledge.
- Pigden, Charles 2009. 'If Not Non-Cognitivism, Then What?' in Pigden, ed., 2009: 80–104.
- , 2010. 'Snare's Puzzle/Hume's Purpose: Non-Cognitivism and What Hume was Really Up To with No-Ought-From-Is', in C. Pigden, ed., *Hume on Is and Ought*. Basingstoke: Palgrave, 169–191.
- , ed., 2009. *Hume on Motivation and Virtue*. Basingstoke: Palgrave Macmillan.
- Radcliffe, Elizabeth 2006. 'Moral Internalism and Moral Cognitivism in Hume's Metaethics', *Synthese* 152: 353–70.
- , 2018. *Hume, Passion, and Action*. Oxford: Oxford University Press.
- Ridge, Michael 2007. 'Ecumenical Expressivism: The Best of Both Worlds?' in R. Shafer-Landau, ed., *Oxford Studies in Metaethics* vol. 2. Oxford: Oxford University Press, 51–76.
- Sayre-McCord, Geoffrey 2008. 'Hume on Practical Morality and Inert Reason', in R. Shafer-Landau, ed., *Oxford Studies in Metaethics* vol. 3. Oxford: Oxford University Press, 299–320.
- Shecaira, Fabio 2011. 'Hume and Non-Cognitivism', *History of Philosophy Quarterly* 28: 267–87.
- Smith, Michael 1994. *The Moral Problem*. Oxford: Blackwell.
- , 2009. 'The Motivation Argument for Non-Cognitivism', in Pigden, ed., 2009: 105–20.
- Snare, Francis 1991. *Morals, Motivation, and Convention: Hume's Influential Doctrines*. Cambridge: Cambridge University Press.
- Sobel, Jordan Howard 2009. *Walls and Vaults: A Natural Science of Morals (Virtue Ethics According to David Hume)*. Hoboken, N.J.: Wiley.
- Sturgeon, Nicholas 2008. 'Hume's Metaethics: Is Hume a Moral Non-cognitivist?' in E. S. Radcliffe, ed., *A Companion to Hume*. Oxford: Wiley-Blackwell, 513–28.





Aleksandra Davidović  
University of Belgrade,  
Faculty of Philosophy,  
Institute for Philosophy of the Department of Philosophy  
alexandra.davidovic@gmail.com

Original Scientific Paper  
UDC 141.4 Xjym Д.

## WHO SPEAKS FOR HUME: HUME'S PRESENCE IN THE *DIALOGUES CONCERNING NATURAL RELIGION*\*

**Abstract:** *One of the reasons for many different and even opposing interpretations of Hume's Dialogues Concerning Natural Religion is the absence of consensus concerning the question of which character in the Dialogues represents Hume. In this paper I argue that taking Philo to be his primary spokesperson provides us with the most consistent reading of the whole work and helps us better understand Hume's religious viewpoint. I first stress the specific dialogue form of Hume's work, which requires us to take into account literary tools such as irony and double-talk when interpreting it. From there I proceed to show why I believe that my hypothesis is better supported than the other two main hypotheses concerning Hume's presence in the Dialogues, the first one being that Cleanthes represents Hume and the other one that none of the characters consistently speaks for Hume but rather that the whole structure of the work does that. Although there is both textual and historical evidence which suggests that Hume favoured Cleanthes, I show that his opinions deviate from Hume's well-known views on important subjects such as scepticism, morality and Christianity, while Philo's opinions on these subjects agree with Hume's almost verbatim. The second hypothesis is proven to be wrong by the fact that Philo actually consistently defends Hume's opinions. Finally, I argue that Philo's understanding of true religion as a philosophical position devoid of any religious import agrees with Hume's religious scepticism.*

**Keywords:** *David Hume, philosophical dialogue, true religion, religious scepticism*

### Introduction

The question of which character in the *Dialogues concerning Natural Religion* represents Hume's own views, and whether this role should be assigned to any one of them, has been widely debated, but still remains undecided. This question is important because identifying Hume's spokesperson in the *Dialogues* could bring us closer to understanding his intentions and goals in writing this work and, in connection to this and maybe more importantly, it could lead us to a better understanding of his own views on religion. It

---

\* I would like to thank an anonymous referee for helpful comments on the previous version of this paper.

is worthy of mention that “[i]n the totality of his work Hume wrote more about religion than about any other single philosophical subject.” (Gaskin 1988, 1) Probably better known than this is the fact that during his life he was notorious for his attitude towards religion, which twice prevented him from obtaining an academic position and almost got him excommunicated from the Church of Scotland. He certainly approached the subject polemically and from many different angles. In both his documented correspondence and in his writings he was highly critical of and even hostile to organised religion, in his *Natural History of Religion* he attempted to explain the causes of religious beliefs exclusively in terms of psychological and sociological factors and throughout his various philosophical writings he contributed to many of the central questions in the philosophy of religion including the reliability of miracles, the immateriality and immortality of the soul, various arguments for God’s existence, God’s natural and moral attributes, the morality of suicide and others. Although none of his arguments is particularly favourable to the religious point of view we are still finding it difficult to decide what they amount to and how Hume’s position concerning religious beliefs should best be characterised.

We could plausibly understand him to be a deist, an atheist, or a sceptic. There are two main reasons for such differing interpretations. The first reason, which I shall discuss only briefly, is that Hume, like every other person engaging critically with religion in the 18<sup>th</sup> century, had to be very careful about what he was saying and writing in order to avoid religious persecution. He couldn’t have voiced some opinions even if he had held them for fear of excommunication, imprisonment or worse. He obviously, and probably intentionally, wasn’t careful enough, which earned him the reputation of an infidel, but it is still true that “in his published writings he had always refrained from marshalling all of his skeptical challenges at once, thereby appearing to leave some kind of refuge for the devout.” (Rasmussen 2018, 188) In his *Treatise on Human Nature* he attacks the causal maxim and the idea of demonstrating the existence of anything *a priori*, thereby undermining the cosmological and the ontological argument for God’s existence, but apparently leaving room for the teleological argument, i.e. the argument from design.<sup>1</sup> In his first *Enquiry* he famously attacks this argument as well, but almost completely

---

1 Hume certainly took advantage of this circumstance when he defended himself from the accusations of maintaining atheism in the *Treatise* in his anonymous *A Letter from a Gentleman to His Friend in Edinburgh*. He replied to his accuser by claiming that the fact that he had denied the principle that whatever begins to exist must have a cause of its existence, thereby undermining Clarke’s cosmological argument for God’s existence, doesn’t mean that he had denied the success of any other argument for God’s existence and he specifically mentioned the teleological argument as one which he, apparently, endorsed: “Wherever I see Order, I infer from Experience that *there*, there hath been Design and Contrivance. And the same Principle which leads me into this Inference, when I contemplate a Building, regular and beautiful in its whole Frame and Structure; the same Principle obliges me to infer an infinitely perfect Architect, from the infinite Art and Contrivance which is displayed in the whole Fabric of the Universe.” (LG 158)

avoids the talk of moral and political effects of religion, and the same is true of all his writings on religion published during his lifetime. However, in the posthumously published *Dialogues concerning Natural Religion*, his greatest and most influential work in the philosophy of religion, Hume carries out the most comprehensive attack on prominent religious doctrines of his time, leaving barely any claim of religious significance intact. Since he knew the *Dialogues* was going to be published posthumously, it might seem that he could have expressed himself more freely there than in any of his other works. However, both how the *Dialogues* is structured and the way some of its parts were written suggest that Hume was concerned for his reputation even after his death and that he took great care to rather carefully and covertly express the most controversial aspects of this work.<sup>2</sup>

The second reason hindering clearer understanding of Hume's religious stance is linked to the first one but specifically concerns the *Dialogues*. The main problem with understanding this work consists in there being many different and even opposing interpretations of it. Since Hume's writing is as clear and precise there as in any of his other works I think this is at least in part a consequence of different interpreters taking different characters to speak for Hume. This was made possible by the infamously ambiguous last part of the *Dialogues*. Had Hume never written it or had he written something like an epilogue to and consistent with the previous eleven parts, I believe that the vast majority of interpreters undoubtedly would have agreed that the sceptic Philo represented his opinions. However, the actual Part 12 of the *Dialogues* contains the so-called "Philo's reversal" in which this character, who previously provided us with what is probably the most comprehensive and powerful critique of the argument from design, professes his "veneration for true religion" and apparently endorses the same argument he had been vehemently criticising thus far. This turn of events at the end of the *Dialogues* confused the interpreters so much that some of them thought that Hume concluded that no religious belief was justifiable while others maintained that he not only accepted the justifiability of some religious beliefs but that he also established that the belief in intelligent design was the most reasonable one. For this reason, I take this to be the main point from which all the different interpretations of Hume's position concerning the existence and nature of God proceeded from.

In what follows I shall argue that, despite these difficulties, we should still take Philo to be Hume's main representative in the *Dialogues*. I will offer reasons as to why I think this hypothesis fares better than the two other hypotheses which have enjoyed some support in both earlier and recent times. My main task, however, will be to show why Philo's apparent change of heart at the end of the *Dialogues* poses no real problem for the interpretation I am defending. This will require me to consider a few different factors that should be taken into account when interpreting the specific type of philosophical

---

<sup>2</sup> More on this in section "The literary character of the *Dialogues*" below.

work such as the *Dialogues*, before I focus on its final part in an attempt to show that, despite how it might seem at first sight, Hume actually remains consistent in his opinions throughout the whole work. I will also briefly touch on the problem of understanding “true religion” in the *Dialogues*, but only insofar as it relates to the problem of identifying Hume’s representative. My hope is that these considerations could make a small contribution to the shaping of a clearer picture of Hume’s views on religion and his most comprehensive work concerning it.<sup>3</sup>

### The characters and the two other hypotheses regarding Hume’s spokesperson

The main theme of the *Dialogues* is answering the question of whether religious beliefs could be rationally founded (i.e. if they could be sufficiently supported by experiential evidence) by way of critically evaluating the teleological argument for God’s existence. This argument is meant to show that, since the order and purpose we find in the world are like the order and purpose present in the products of human artifice, they also must be the result of the creation of an intelligent designer. Hume decides to set aside the question of whether we can prove God’s existence and to consider only whether it is possible to reasonably conclude something about his nature, but by the end of the *Dialogues* it is clear that these two questions cannot be completely separated one from another.

Three characters are taking part in this discussion: the experimental theist Cleanthes, who thinks that religious beliefs have solid experiential support and who is the main proponent of the argument from design, the philosophical sceptic Philo, who thinks that such subjects fall outside the scope of human reason and knowledge, and the traditional orthodox theist Demea, who believes that the only way to gain knowledge of God’s nature is through revelation. Cleanthes’ pupil Pamphilus is also present during the conversation and the *Dialogues* we are reading is Pamphilus’ recounting of it in a letter to his friend Hermippus, who was rather intrigued by the first

3 Although identifying Hume’s spokesperson would certainly enable us to gain a better understanding of the *Dialogues* and possibly of Hume’s final position concerning religion as well, I would like to note that I don’t think it would help us reach a definitive interpretation of this work. I highly doubt that any such interpretation is actually possible. As I briefly explain in the third section of the paper, I believe that Hume wanted different people to read and take something away from this work, which would not be possible were there only one way to understand its meaning. This should not be taken as undermining the thesis that Philo is Hume’s main representative, but it does mean that, even if we grant this thesis to be true or at least the most plausible one, we are still facing some difficulties when it comes to understanding what exactly Hume wanted to convey through Philo. However, I believe that every interpretation will be similarly limited, given the inherent and intentional ambiguities of the *Dialogues’* text, and that the one I am defending is still best supported by various evidence I shall present in the course of this paper.

account of the discussion Pamphilus had witnessed and who had asked for a more detailed report of it. These might sound like minor details, but they will become important later on.

Throughout the *Dialogues* we see Philo constantly challenging and opposing Cleanthes' endeavours to establish the design hypothesis and he seems to be doing this with great success. He argues against our ability to infer from experience both God's natural and moral attributes and in response to his critiques laid out in Part 10<sup>4</sup> Cleanthes even suggests that God is not infinitely but only finitely perfect. However, in the last section of the *Dialogues*, in a *tête-à-tête* with Cleanthes Philo suddenly confesses that, despite his "love of singular arguments, no one has a deeper sense of religion impressed on his mind, or pays more profound adoration to the divine Being, as he discovers himself to reason, in the inexplicable contrivance and artifice of nature". (DNR 12.2; KS 214) He then speaks very laudably about the argument from design and acts as if it were impossible for him to doubt it, and even produces several examples of the evidence of design in the world. He goes so far as to express his "veneration for true religion" (DNR 12.9; KS 219) although he doesn't say what he means by "true religion". Finally, at the very end of the *Dialogues* Philo explicitly accepts sceptical fideism by saying that "to be a philosophical Sceptic is, in a man of letters, the first and most essential step towards being a sound believing Christian." (DNR 12.33; KS 228) It is this part of the *Dialogues* and these claims that have usually been described as containing and expressing "Philo's reversal" and which have generated such controversy among the interpreters.

Of the five characters in total who are mentioned in the *Dialogues*, three are participants in the philosophical discussion. Nowhere in the literature have I been able to find anyone who thinks that Demea might be Hume's primary spokesperson,<sup>5</sup> and there are good reasons for this. Demea is a rigid orthodox theist and occasionally a mystic. He thinks that God's nature is inscrutable – "altogether incomprehensible and unknown to us" (DNR 2.1; KS 141) which is why he opposes the idea of deciding anything concerning his nature on the basis of experience. When pressed to engage with the

4 Philo's main challenge in Part 10 consists in asking these questions about the being whose moral attributes Cleanthes wishes to establish: "Is he willing to prevent evil, but not able? then is he impotent. Is he able, but not willing? then is he malevolent. Is he both able and willing? whence then is evil?" (DNR 10.25; KS 198) At the beginning of Part 11 Cleanthes attempts to answer the challenge by claiming that "supposing the Author of Nature to be finitely perfect, though far exceeding mankind; a satisfactory account may then be given of natural and moral evil, and every untoward phenomenon be explained and adjusted." (DNR 11.1; KS 203)

5 The qualification "primary" is important because neither I nor the other authors writing about this question wish to say that Demea never speaks for Hume. However, when he does that, Philo and (in most cases) Cleanthes agree with him, while there is no single instance of him defending Hume's known views against the opposing views of those two characters.

philosophical arguments he proposes “that simple and sublime argument *a priori*” (DNR 9.1; KS 188), a version of the cosmological argument for God’s existence very similar to the one that Samuel Clarke<sup>6</sup> defended and which Hume had already criticised.<sup>7</sup> This clearly makes Demea a bad candidate for Hume’s main representative although it doesn’t mean that he is not an important character in his own right.<sup>8</sup> However, ruling him out as Hume’s primary spokesperson leaves us with Philo and Cleanthes as main contenders.

Most authors today take Philo to be the obvious choice between the two or opt for the third interpretative solution which I shall mention shortly. However, before Kemp Smith’s analysis of the *Dialogues* (in 1947) the most popular hypothesis was that Cleanthes was Hume’s spokesman, and even after he offered his interpretation, according to which “Philo, from start to finish, represents Hume” (Kemp Smith 1947, 59), Harward claimed that both Philo and Cleanthes represented Hume (Harward 1975). On the other hand, and perhaps surprisingly considering that Pamphilus and Hermippus are clearly minor characters, Wieand claimed that Pamphilus should be taken as Hume’s spokesperson because he is the narrator of the work who both introduces the characters and their discussion and passes the final judgement about the winner of the debate (Wieand 1985, 33–35). I believe this is the only suggestion of that kind but even the authors who wouldn’t accept Pamphilus to be Hume’s representative throughout the *Dialogues* might still think that he is speaking for Hume in the moments when he is passing those judgements. Because Pamphilus speaks very little and when he does he always agrees with Cleanthes or supports his opinions (which is only natural, considering that he is his pupil), I shall deal with the suggestions that Hume might be represented by Cleanthes and by Pamphilus together.

---

6 It is sometimes suggested that Hume modelled Demea on Samuel Clarke (notably in Mossner 1936, as well as in Kemp Smith’s Introduction to the *Dialogues*), perhaps because he is the one who presents the cosmological argument in the *Dialogues*, but this is not very likely because Demea’s mysticism doesn’t align well with Clarke’s rationalism and Clarke was certainly a better philosopher than Demea.

7 In the *Treatise* (T 1.3.3) Hume discusses “a general maxim in philosophy, *that whatever begins to exist, must have a cause of existence*” (the maxim is a version of the famous principle of sufficient reason) and during this discussion he examines and deems question-begging Clarke’s proof of the maxim. This proof and the maxim itself figure prominently in Clarke’s cosmological argument in his *A Demonstration of the Being and Attributes of God* which is why Hume was accused of denying this argument. See fn. 1.

8 While he was virtually ignored or thought wholly insignificant before, the past few decades saw the emergence of new interest in Demea’s role in the *Dialogues*. Thus Stove (1978), Stahl (1984), and Dye (1989) all discuss, more or less critically, Demea’s arguments in Part 9. On the other hand, Vink (1986) and Dye (1992) recognize the importance of Demea’s departure at the end of Part 11 for the overall direction of Part 12 of the *Dialogues*, while Olsewsky claims that even before his departure “Demea has a crucial and central role in the dramatic triangularity that gives impetus to the dialogue.” (Olsewsky 2003, 474–75) However, none of these authors suggests that Demea might be Hume’s main representative.

The other hypothesis that I would like to consider and one that I feel has more substantial support is that no one in particular speaks for Hume in the *Dialogues*. This could be understood as the suggestion that more than one character speaks for Hume or that none of them speak for him but the main point is that, as Fogelin put it, “the entire dialectical structure of the *Dialogues* speaks for Hume.” (Fogelin 2017, 101) These two alternatives are often combined to express the claim that, although all of the characters sometimes speak for Hume, none of them does so consistently but rather the *Dialogues* as a whole, when properly understood, reveal his genuine opinions. Thus Brick thinks that the fundamental mistake in interpreting the *Dialogues*, which leads to many other mistakes, consists in believing that one character speaks for Hume (Bricke 1975, 3). Foley defends the same view and claims that the characters in the *Dialogues* are rather inconsistent and that we can take Hume to be consistent only if we accept that none of them is representing his opinions (Foley 2006, 84). I find this interpretation rather plausible and valuable because it stresses the need to think about the *Dialogues* in its entirety and to also take into account the fact that it is a special kind of literary work. What I do not agree with is the claim that none of the characters consistently represents Hume’s opinions. In what follows I will first draw attention to some of the literary aspects of the *Dialogues* that I think are necessary to take into consideration when interpreting the work and then I will proceed to show why Cleanthes or Pamphilus are not the right candidates for Hume’s representatives. Finally, my main argument against the hypothesis that the whole work speaks for Hume will be that Philo is actually consistently defending his views and that we can gain a better understanding of Hume’s own views if we consider the *Dialogues* in this light rather than if we take none of the characters to have this role.

### The literary character of the *Dialogues*

The *Dialogues* is Hume’s posthumously published work and we don’t know much about his process of writing it. What we know for sure is that Hume worked on it on at least three different occasions over a period of twenty-five years and that he spent what he knew to be the last months of his life revising this work. We also know that many of his friends had endeavoured and finally managed to persuade him not to publish the *Dialogues* earlier, but that he took great care to make sure it would be published after his death.<sup>9</sup> It

9 It turned out that this would be more difficult to achieve than Hume had expected. He first tried very hard to persuade his good friend Adam Smith to publish it but with no success. Smith wrote to William Strahan that he “could have wished [that the *Dialogues*] had remained in manuscript to be communicated only to a few people” (HL2, 453) and he clearly didn’t want his name to be in any way connected to this work. Strahan, who was Hume’s publisher, initially agreed to publish it but after Hume’s death he obviously deemed it improper and refrained from complying with his wishes. Finally, Hume’s

was obviously important to Hume that his work sees the light of day and he seemed to be satisfied with the final version of it. In a letter to Adam Smith, written ten days before his death, he says that after the newest revisions he found “that nothing can be more *cautiously* and more *artfully* written.” (HL2, 334, my emphasis) We could understand the word “artfully” to mean “with artistic skill”, but I believe that we should actually take it to mean “skilfully” or maybe “cunningly”. The fact that Hume also says that the *Dialogues* is “cautiously” written suggests that he was well aware of its contents being controversial. He wrote the letter to tell Smith that he shouldn’t have scrupled to publish the *Dialogues* (which Smith had done), so we have reasons to suppose that he wanted to tell him that whatever was objectionable in the work was at least expressed very carefully.

Why did Hume feel the need to be so cautious when he knew the *Dialogues* were going to be published posthumously? I believe there are several reasons for this. On the one hand, he wanted to make sure that his work would actually be published, and he realised how difficult that would be if the *Dialogues* were thought to further atheism. On the other hand, I don’t think it was Hume’s intention at all to give ready answers in his *Dialogues*. We know that he revised and expanded Part 12 shortly before he died<sup>10</sup> and it is precisely that part which prevents us from readily answering questions concerning his representative and his religious views in the *Dialogues*. The final version of the text made several interpretations at least plausible. This meant that different thinkers could gain something from engaging with his work without outright condemning it for atheism, even if it was still undoubtedly perceived as prevalently anti-religious. I also believe, but I cannot argue for that point in more detail here, that Hume didn’t mean to conceal either his atheism or his deism in the *Dialogues*, but rather to really subtly argue for scepticism. If this were correct it would explain why it is so difficult to find any definite answers in this work.

The *Dialogues* is not only a philosophical work, but a work of literature at the same time. It is, as its title says, written in dialogue form, a kind of literary form which Hume didn’t use very often but which he seems to have considered especially apt for writing about the subjects he writes about in

---

nephew David Hume Younger saw it through the press in 1779, three years after his uncle’s death and according to his wish. Hume was obviously uncertain that any of his friends would have the courage to publish the *Dialogues* and he wrote to Strahan that if, for whatever reason, he did not publish the work within two and a half years of Hume’s death then “the Property shall return to my Nephew, David, whose Duty, in publishing them as the last Request of his Uncle, must be approved of by all the World.” (*ibid.*) For a more detailed story about the publication of the *Dialogues* I would like to direct readers to Chapter 10 of Rasmussen’s book, as well as to the Appendix M of HL2.

10 Stewart has shown that the only paragraph added in 1776 was the longest paragraph in Part 12, where Philo argues that the dispute between theists and atheists is purely verbal (Stewart 2000, 303). I discuss this paragraph briefly in the penultimate section of the paper.



the *Dialogues*. Another one of his writings that features this literary form is Section 11 of the first *Enquiry*, in which he also examines the argument from design. I already noted that the structure of the *Dialogues* is very peculiar<sup>11</sup> – we as readers share Hermippus’ position and hear about the conversation between three interlocutors from Pamphilus, who is the pupil and protégé of Cleanthes. He doesn’t hesitate to express his agreement with his teacher and we may rightly wonder how reliable a narrator he is. In addition to this, Hume’s *Dialogues* differs from the majority of other philosophical dialogues, both ancient and modern, precisely in the fact that there is not one dominant character of whom we can say with certainty that this person represents the author’s views.

From the introductory part of the *Dialogues*, through the words of Pamphilus, we learn that Hume gave much thought to the specificities, both advantages and disadvantages, of this literary form, especially when philosophers make use of it. Pamphilus informs us that in a dialogue “the order, brevity, and precision” typical for a philosophical exposition are sacrificed to preparations, transitions and a mix of various topics thrown in to make the conversation sound more natural (DNR 0.1; KS 127). However, there are certain subjects “to which dialogue-writing is peculiarly adapted, and where it is still preferable to the direct and simple method of composition.” (DNR 0.2; KS 127) Such subjects are at the same time obvious and important, but also both obscure and uncertain, and Pamphilus remarks that “these circumstances are all to be found in the subject of natural religion” (DNR 0.5; KS 128). He also explains that the dialogue-writer who wants to carry “the dispute in the natural spirit of good company” tries to preserve “a proper balance among the speakers” (DNR 0.1; KS 127), which is one of the reasons why it is so difficult to identify Hume’s spokesperson, but he never says that there can be no such person.

So we are warned right from the beginning about what we should and shouldn’t expect to find in Hume’s *Dialogues*, which Vink helpfully summarizes in this way: “[w]e may expect the relations among the participants in the discussion to be of importance; a number of problems to be discussed, not all of them as thorough as perhaps should be; a certain balance among the speakers to be aimed at deliberately; some passages to function as preparations and transitions; order, brevity and precision not to be the main concern.” (Vink 1986, 389) That is why it is surprising to see that Bricke, who justly claims that the right way of approaching the *Dialogues* is “one

---

11 The dialogue in the *Enquiry* also has a rather unusual structure – Hume is recounting to the readers the conversation he recently had with a friend “who loves sceptical paradoxes” and in this conversation they enacted Epicurus’ defence before the Athenian people, where Hume represented the Athenian people while his friend, in the role of Epicurus, argued against the teleological argument. Hume also took care to mention right at the beginning that this friend of his “advanced many principles, of which I can by no means approve” (EHU 11.1; SBN 132).

which puts proper stress on Hume's literary objectives in their composition" (Bricke 1975, 3) also thinks "that neither Philo nor Cleanthes [...] could be Hume's spokesman, for each is often unHumean in his views, and neither maintains a reasonably clear, well-argued, self-consistent position in the course of the *Dialogues*" (*ibid.*). I will address the claim that both characters are unHumean below, but I would like to point out here that, if we really pay attention to Hume's literary objectives, then the fact that his characters aren't always well-argued or wholly consistent shouldn't come as a surprise. Hume already informed us that he didn't create his dialogue or his characters to be that way and Bricke's contention is not a valid objection to any of them being Hume's spokesperson.

Finally, it is well-known that Hume uses irony in the *Dialogues*, especially through the character of Philo. This further complicates the interpretation of this work because the irony is sometimes very obvious (as when Philo represents himself as Demea's ally) but at other times we could be missing it and, in consequence, wrongly interpreting certain parts of the discussion. On the other side, there is the opposite danger of thinking that Hume is being ironic when he actually isn't or of trying to explain away any difficult passages by ascribing them to Hume's or Philo's irony. Because of all of this we should be very careful when we are attributing certain statements to either one of the characters or to the author and it is important to remember at all times the aforementioned factors we ought to take into account when interpreting the *Dialogues*.

## Against the view that Cleanthes speaks for Hume

At the beginning of the *Dialogues* we are also told about the "remarkable contrast" in the characters of the three interlocutors and we learn that, upon Pamphilus' first ("imperfect") account of the conversation he had witnessed, Hermippus "opposed the accurate philosophical turn of Cleanthes to the careless scepticism of Philo" and "compared either of their dispositions with the rigid inflexible orthodoxy of Demea" (DNR 0.6; KS 129), while at the very end of the *Dialogues*, after seriously considering the whole discussion, Pamphilus pronounces "that Philo's principles are more probable than Demea's; but that those of Cleanthes approach still nearer to the truth." (DNR 12.34; KS 228) Here we seem to have some direct evidence for the interpretation that Cleanthes represents Hume. In addition to this, in a letter to his friend Gilbert Elliot Hume writes: "You would perceive by the Sample I have given you, that I make Cleanthes the Hero of the Dialogue. Whatever you can think of, to strengthen that Side of the Argument, will be most acceptable to me." (HL1, 153–154) Taken at face value, this would seem to confirm Pamphilus' verdict. However, apart from our having reasons to suspect Pamphilus' impartiality, what Hume says in the letter seems strange. We might ask why would he need his friend's help in strengthening Cleanthes' argument if he himself was a genuine proponent of that argument. Besides that, he didn't

really say that Cleanthes *was* the hero of the *Dialogues*, but only that *he made him so*. He continues the letter: “Any Propensity you imagine I have to the other Side, crept in upon me against my Will: And tis not long ago that I burned an old Manuscript Book, wrote before I was twenty; which contained, Page after Page, the gradual Progress of my Thoughts on that head.” (HL1, 154) Despite the proclamation that it crept in upon him against his will, this implies that Hume himself felt he had more affinity for Philo’s than for Cleanthes’ views. If this was the case it would make sense that he needed help in strengthening Cleanthes’ argument, whatever his reasons for wishing it might have been. Finally, “the old manuscript book” Hume mentioned in the letter is probably something he wrote while he was writing the *Treatise*, and maybe even something he had intended to include in the *Treatise* but later refrained from it.<sup>12</sup> In any case, it must have contained sceptical thoughts on religion, since Hume here claims that it was in line with Philo’s side of the argument.

Although Hume’s *Dialogues* differ from the majority of other well-known philosophical dialogues, we find one important exception. There is actually a striking similarity between Hume’s *Dialogues* and Cicero’s *De natura deorum* which should not be overlooked because it potentially provides us with very helpful insights into Hume’s structuring the *Dialogues* precisely the way he did.<sup>13</sup> In Cicero’s dialogue there are also three main participants: Velleius – the Epicurean, Balbus – the Stoic, and Cotta – the Academic Sceptic. Cicero himself is present mainly as a narrator and listener, in a similar way in which Pamphilus is present in Hume’s dialogue, but historical Cicero obviously endorsed the views advanced by Cotta, who clearly wins the debate. However, *De Natura Deorum* ends with Cicero saying: “Cotta’s argument seemed to Velleius to be more truthful, but in my eyes Balbus’ case seemed to come more closely to a semblance of the truth.” (Cicero 1997, 146) We know that Hume greatly admired Cicero and his style of writing so it is not surprising that he took inspiration from Cicero’s dialogue’s ending.<sup>14</sup> If we have this in

---

12 We already know that Hume wrote the first draft of his essay “On miracles” (Section 10 of the first *Enquiry*) well before he published the *Treatise*, because he wrote to Henry Home in 1737: “[I] enclose some *Reasonings concerning Miracles*, which I once thought of publishing with the rest, but which I am afraid will give too much offence” (HL1, 24) and in the same letter he informed Home that he was at the time castrating his work, “that is, cutting off its nobler parts, that is, endeavouring it shall give as little offence as possible” before he could show it to Bishop Butler (HL1, 25). It is possible that Hume wrote another piece on religion at this time which he decided not to publish with the rest.

13 Cicero’s influence on the *Dialogues* has been widely recognised, particularly in Price 1964, Battersby 1979, and Fosl 1994.

14 Hume took inspiration from Cicero for more than just the ending of the *Dialogues*. In Cicero’s dialogue Velleius says that Philo (talking about Philo of Larissa, an Academic Sceptic) was both Cicero’s and Cotta’s teacher: “both of you have been taught by the same teacher Philo to know nothing.” (Cicero 1997, 9) Balbus, on the other hand, mentions that Cleanthes belongs to their (Stoic) school (Cicero 1997, 52) – meaning Cleanthes of

mind and the fact that Cleanthes clearly wasn't the winner of the debate in the *Dialogues*, then the fact that Pamphilus considers him to be the winner and that Hume names him the hero of the *Dialogues* doesn't provide us with good reasons to think that Cleanthes represents Hume's views.

In the previously cited letter to Elliot Hume also writes the following:

“I have often thought that the best way of composing a dialogue would be for two persons that are of different opinions about any question of importance, to write alternately the different parts of the discourse, & reply to each other. By this means, that vulgar error would be avoided, of putting nothing but nonsense into the mouth of the adversary: And at the same time, a variety of character & genius being upheld, would make the whole look more natural & unaffected. Had it been my good fortune to live near you, *I should have taken on me the character of Philo in the dialogue, which you'll own I could have supported naturally enough*: And you would not have been averse to that of Cleanthes.” (HL1, 154, my emphasis)

Here we can see again how much Hume wanted his dialogue to give the impression of being natural and his characters to feel genuine. He obviously disliked dialogues of a Socratic kind and he made a conscious effort not to have one character dominate the conversation the whole time. However, the fact that he said that he would have left the role of Cleanthes to Elliot, while he himself could have supported Philo “naturally enough” also implies that the intended hero of the *Dialogues* is not the one with the views closest to his own. Therefore, neither Pamphilus' verdict at the end of the *Dialogues* nor Hume's letter to Elliot provide good evidence for the hypothesis that Cleanthes represents Hume in the *Dialogues*.

## Philo's scepticism

What about Hermippus' opposition<sup>15</sup> of “the accurate philosophical turn of Cleanthes to the careless scepticism of Philo” (DNR 0.6; KS 129)?

Assos, a Greek Stoic philosopher and successor to Zeno of Citium as the second head of the Stoic school in Athens.

15 My referring to *Hermippus'* comparing and contrasting the characters here might seem strange because he himself is barely a character in the *Dialogues* and he never utters a word throughout the book. However, Pamphilus explicitly says in the Introduction that Hermippus' expectations were raised by “the remarkable contrast” in the interlocutors' characters and that Hermippus opposed different philosophical approaches of Cleanthes, Philo and Demea. He also informs us that Hermippus shared these opinions with him after having read Pamphilus' first and imperfect account of the conversation between the three main characters. This exchange of letters took place before Pamphilus' narrating the *Dialogues* to Hermippus and to us, and everything we know of Hermippus' opinions comes from Pamphilus who never informs us whether Hermippus' views were changed after the second, presumably less imperfect, account. Without delving into the question of why Hume introduced the character of Hermippus at all while at the same time

Philo's scepticism is sometimes characterised as excessive. At the beginning of the *Dialogues* Cleanthes accuses him of advocating Pyrrhonism, but Philo explains that his scepticism is of the kind that Hume openly accepts and calls moderate or mitigated scepticism. He doesn't use that term but he describes his scepticism in much the same way and using almost the same words that Hume uses in the first *Enquiry*. When Cleanthes first challenges the tenability of Philo's sceptical position he is being rather ironical: "Whether your scepticism be as absolute and sincere as you pretend, we shall learn by and by, when the company breaks up: We shall then see, whether you go out at the door or the window; and whether you really doubt, if your body has gravity, or can be injured by its fall [...]" (DNR 1.5; KS 132). He then proceeds in a more serious tone to show that human beings are incapable of enduring total scepticism. His main point is that, even though a man may momentarily, after intensely reflecting on the imperfections of our faculties and the groundlessness of our beliefs, "entirely renounce all belief and opinion, it is impossible for him to persevere in this total scepticism, or make it appear in his conduct for a few hours." (*ibid.*) He may be able to follow these sceptical principles while he is engaging with philosophical arguments, but as soon as he turns to the more mundane things "the bent of his mind relaxes" and "the philosopher sinks by degrees into the plebeian." (DNR 1.7; KS 133) According to him sceptical philosophy simply cannot have any lasting influence.

While this critique of excessive scepticism is entirely Humean, and mirrors Hume's earlier critiques of what he calls Pyrrhonism in the *Treatise* and in the first *Enquiry*, Philo's answer is not less Humean. He agrees with Cleanthes that, as much as someone may be sceptical in philosophy, "he must act, I own, and live, and converse like other men; and for this conduct he is not obliged to give any other reason, than the absolute necessity he lies under of so doing." (DNR 1.9; KS 134) In other words, although philosophy teaches us that we are by no means able to justify even our most fundamental beliefs and that our faculties are constantly leading us to erroneous conclusions, we continue to have those beliefs and use those faculties because we cannot help it – "[n]ature, by an absolute and uncontrollable necessity has determined us to judge as well as to breathe and feel" (T 1.4.1.7; SBN 183), as Hume himself says in the *Treatise*. However, this doesn't mean that scepticism cannot have any lasting influence. Contrary to that, Philo claims that a man who "has accustomed himself to sceptical considerations" won't entirely forget them, but he will, both in his philosophy and in his everyday conduct, differ from those who have never experienced such sceptical doubts. (DNR 1.9; KS 134) How will he differ from them? He will continue to philosophise because he finds pleasure in it,<sup>16</sup> however, he will confine his enquiries to the subjects

---

distancing the readers from him as much as possible, I will refer to Hermippus' opinions wherever Pamphilus refers to them in such way.

16 This is exactly what Hume cites as his reason to continue doing philosophy after serious sceptical crisis at the end of the Book I of the *Treatise*. He describes the curiosity about

that are within the bounds of our experience and concerning which we can “make continual advances in forming more general principles of conduct and reasoning”, but he won’t carry his speculations much farther (DNR 1.9– 1.10; KS 134– 135). This is quite similar to Hume’s position at the end of the first *Enquiry*, where he claims that mitigated scepticism or academical philosophy, “which may be both durable and useful,” is in part “the result of this Pyrrhonism, or *excessive* scepticism, when its undistinguished doubts are, in some measure, corrected by common sense and reflection.” (EHU 12.24; SBN 161) He also explicitly says that one aspect of this philosophy, which is the natural result of Pyrrhonism, is “the limitation of our enquiries to such subjects as are best adapted to the narrow capacity of human understanding.” (EHU 12.25; SBN 162) He ends these considerations with the following words:

“Those who have a propensity to philosophy, will still continue their researches [...] But they will never be tempted to go beyond common life, so long as they consider the imperfection of those faculties which they employ, their narrow reach, and their inaccurate operations. While we cannot give a satisfactory reason, why we believe, after a thousand experiments, that a stone will fall, or fire burn; can we ever satisfy ourselves concerning any determination, which we may form, with regard to the origin of worlds, and the situation of nature, from, and to eternity?” (EHU 12.25; SBN 162)

And in the *Dialogues* Philo says:

“Let us become thoroughly sensible of the weakness, blindness, and narrow limits of human reason: Let us duly consider its uncertainty and endless contrarieties, even in subjects of common life and practice [...] When the coherence of the parts of a stone, or even that composition of parts, which renders it extended; when these familiar objects, I say, are so inexplicable, and contain circumstances so repugnant and contradictory; with what assurance can we decide concerning the origin of worlds, or trace their history from eternity to eternity?” (DNR 1.3; KS 132)

I believe that this strongly suggests that Hume and Philo at least share the same sceptical outlook. To return now to Hermippus’ characterisation, we don’t have to understand “careless” to mean “reckless”. We could understand it to mean “care-free”, as Lorkowski suggests (Lorkowski 2016, 260), since this seems to be the meaning in which Hume used this word in the concluding section of the *Treatise* when he wrote that “[t]he conduct of a man, who studies philosophy in this careless manner, is more truly sceptical than that of one,

---

philosophical questions and the ambition to contribute to the instruction of men naturally arising in him and declares: “should I endeavour to banish them, by attaching myself to any other business or diversion, I *feel* I should be a loser in point of pleasure; and this is the origin of my philosophy.” (T 1.4.7.12, SBN 271).

who feeling in himself an inclination to it, is yet so overwhelmed with doubts and scruples, as totally to reject it.” (T 1.4.7.14; SBN 273) On the other hand, when Hermippus observed that Cleanthes’ philosophical turn was accurate, he didn’t necessarily intend to say that his opinions or conclusions were such as well. Hume would certainly characterise Cleanthes’ use of experimental method in philosophy of religion as accurate, but I don’t think that he would agree with many of his philosophical claims.

### **Against the view that the whole structure of the *Dialogues* speaks for Hume**

Thus far I have only discussed the hypothesis that Cleanthes or Pamphilus represent Hume in the *Dialogues*, but I’ve hardly said anything about the hypothesis that, while all characters speak for Hume, neither one of them does that consistently. However, if I manage to show that Cleanthes doesn’t defend Hume’s opinions at some important moments and that Philo does that at all times, I will have sufficiently argued against both hypotheses. In other words, if I manage to show that Philo consistently defends Hume’s views, then I have *ipso facto* shown that Cleanthes doesn’t and that it is not true that none of the characters is Hume’s primary spokesman. Before I proceed to attempt to do this, I would like to clarify a few things with regard to this hypothesis.

The claim that the *Dialogues* as a whole speak for Hume is trivially correct because he is the author of this work and it reflects in its entirety his process of thinking about natural theology. Saying that “the entire dialectical structure” of the book speaks for Hume is even more accurate because it underlines his struggles in looking for answers to some of the questions he discussed. In spite of that, I don’t see good reasons to favour this hypothesis over the one that Philo primarily speaks for Hume. It is important to understand that, in claiming that Philo represents Hume in the *Dialogues*, I am not claiming that other characters never express any of his opinions. We have already seen that we might plausibly suppose that Pamphilus speaks for Hume when he discusses the peculiarities of the philosophical dialogues at the beginning of the book. There is also an overall consensus among the interpreters that Cleanthes speaks for Hume when he is attacking the ontological and the cosmological arguments in Part 9 of the *Dialogues*, but it is at the same time clear (and Hume makes sure that Cleanthes mentions this in D 9.4; KS 189) that Philo agrees with him. However, “[t]he interpretive challenge is to deduce Hume’s final position in the less obvious cases, i.e. in the points over which characters intelligently disagree.” (Lorkowski 2016, 256) Although both Cleanthes and – to a lesser extent – Demea make references to Humean ideas and sometimes even use Hume’s well-known phrases to do so, in such cases Philo always agrees with them. On the other hand, I believe that it can be shown that Philo is the one who advances Hume’s views when the others

attack them, that he speaks for him most often and most consistently and that his views on religion agree with Hume's views on it that we are familiar with from his other works, from which Cleanthes' and Demea's views clearly differ. That is why I claim that he is not Hume's *only* but, rather, his *primary* representative.

There is also one subsidiary textual evidence against the hypothesis that none of the characters consistently speaks for Hume, which Lorkowski was the first, as far as I know, to take notice of (Lorkowski 2016, 264). In a letter from 1753, written to one of his anonymous reviewers, Hume makes the following complaint: "But you impute to me both the sentiments of the Sceptic and the sentiments of his antagonist, which I can never admit of. *In every Dialogue, no more than one person can be supposed to represent the author.*" (HL1, 173, my emphasis) The dialogue in question is the one appended to his *Enquiry concerning the Principles of Morals* but it is clear from his wording that what Hume says here is meant to apply to *every* philosophical dialogue. Furthermore, we know that he had already written a considerable portion of the *Dialogues*, if not the whole work, by this time, since this letter is dated two years after the aforementioned letter he wrote to Elliot. Therefore, this evidence supports the hypothesis that Hume also thought that only one character should be supposed to represent the author in his *Dialogues*. In what follows I hope to show this to be true even regardless of this evidence by showing that no one but Philo consistently speaks for him throughout the *Dialogues*.

### "Philo's reversal" in DNR 12

If everything that I endeavoured to establish thus far is accepted, then the biggest obstacle to identifying Hume with Philo is Philo's alleged change of opinion in Part 12 of the *Dialogues*. This part opens with Philo suddenly confessing to Cleanthes, after Demea had left the company, that despite his "love of singular arguments, no one has a deeper sense of religion impressed on his mind, or pays more profound adoration to the divine Being, as he discovers himself to reason, in the inexplicable contrivance and artifice of nature". (DNR 12.2; KS 214) Furthermore, he claims that "a purpose, an intention, a design strikes everywhere the most careless, the most stupid thinker" (*ibid.*) and that "all the sciences almost lead us insensibly to acknowledge a first intelligent Author." (DNR 12.2; KS 215) In support of this he cites Galen's discoveries concerning the structure of the human body. If a man has above six hundred different muscles, then "whoever duly considers these, will find, that in each of them Nature must have adjusted at least ten different circumstances" so that in the muscles alone "above 6000 several views and intentions must have been formed and executed." (*ibid.*) If we also include in these calculations bones, skin, ligaments, vessels, glandules, and other parts of the body, then, Philo poses a rhetorical question, "how must



our astonishment rise upon us, in proportion to the number and intricacy of the parts so artificially adjusted?” and “to what pitch of pertinacious obstinacy must a philosopher in this age have attained, who can now doubt of a Supreme Intelligence?” (*ibid.*)

Something here already seems a little strange and makes us wonder whether Philo is being ironical or whether he really believes that all these instances of “artificially adjusted” body parts are proofs of intelligent design. If the latter is the case, then why did he not mention it before and why did he try so hard to show that Cleanthes’ examples of instances of design in the world are incapable of establishing the desired conclusion? However, what he says next makes us doubt his honesty even more. Supposing there were an imperceptible God, he asks, “were it possible for him to give stronger proofs of his existence, than what appear on the whole face of Nature?” and he appears to maintain that such a God could do no better “but copy the present economy of things; render many of his artifices so plain, that no stupidity could mistake them; afford glimpses of still greater artifices [...] and conceal altogether a great many from such imperfect creatures?” (DNR 12.4; KS 215–216) He indeed seems to be endorsing the argument from design and he even adds that, “as the works of Nature have a much greater analogy to the effects of our art and contrivance, than to those of our benevolence and justice”, this gives us reasons to conclude that God’s natural attributes resemble man’s more than his moral attributes, which can only mean that man’s moral attributes are more defective since God must be “absolutely and entirely perfect.” (DNR 12.8; KS 219) If we’ve ever read the previous eleven parts of the *Dialogues* we must be extremely confused at this point because there Philo argued over and over again that the argument from design cannot succeed in establishing either God’s natural or his moral attributes, and besides this, I have already mentioned that he made Cleanthes accept in Part 11 that God could only be finitely perfect. Why is he now contradicting himself? He doesn’t give any reasons for suddenly reversing his earlier opinions. He doesn’t even acknowledge that there has been such a reversal of his views and he certainly never refutes any of his previous arguments.

This has led some authors to suggest that there might be some irrational basis for theism in our nature that we simply cannot avoid. I shall only briefly comment on one such view, similarly expressed by Lorkowski (Lorkowski 2014; 2016) and Fogelin (Fogelin 2017), according to which in Part 12 of the *Dialogues* Philo doesn’t endorse the argument he had been refuting thus far, but he actually assents to what is called the irregular teleological argument. In Part 3 of the *Dialogues* Hume draws a difference between the regular and the irregular teleological argument, where by the “regular argument” he means the argument from design that is based on analogical or abductive reasoning, and by the “irregular argument” he means Cleanthes’ assertion that the consideration of purposefulness in nature (the curious adaptation of means to ends) makes the idea of a contriver “immediately flow in upon you

with a force like that of sensation.” (DNR 3.7; 154) Cleanthes in Part 3 claims that even a sceptic must “adhere to common sense and the plain instincts of nature; and to assent, where-ever any reasons strike him with so full a force, that he cannot, without the greatest violence, prevent it” (*ibid.*) Philo didn’t respond to this argument and Pamphilus noticed that he was a little embarrassed and confused upon hearing it (DNR 3.10; KS 155). Lorkowski and Fogelin believe that he remained silent because he couldn’t find a way to oppose this argument and that Hume here wanted to say that religious beliefs are natural in a way. Fogelin describes this sort of belief as “immediate and, when it occurs, irresistible.” (Fogelin 2017, 96) Lorkowski admits that religious beliefs are not natural in the same way beliefs in causation are, because they are neither universal nor necessary for our survival, but he thinks they are still natural in some weaker sense because they are the consequence of “our universal, instinctual propensity to believe that the order of the universe is intended. It is not an inference but is akin to sensation, making the belief non-inferential and therefore natural.” (Lorkowski 2016, 255)

Without entering into discussion concerning the nature of religious belief,<sup>17</sup> I would like to observe that this suggestion, even if we grant that religious beliefs could be in some way natural, doesn’t seem to be a very plausible explanation of “Philo’s reversal”. Firstly, if this was really Hume’s main conclusion concerning religious beliefs, then he did a very lousy job of presenting it to the readers. If Philo was indeed convinced by Cleanthes’ irregular argument in Part 3, then the discussion should have ended then and there or, if it was continuing only in order to show that the belief in an intelligent author cannot be rationally justified, although it was established that it was natural and irresistible, this could and probably would have been made sufficiently clear. Secondly, and more importantly, there is no evidence in Part 12 that an “instinctual propensity to believe that the order of the universe is intended” was what caused Philo to express his “adoration” of the divine being. Furthermore, he explicitly claims that this divine being “discovers himself to reason” which implies that his belief in God is not the result of something “akin to sensation”. He simply doesn’t seem to be assenting to the irregular argument, notwithstanding his claim that a purpose and an intention strikes everywhere “even the most stupid thinker”, which sounds more ironical than anything else, since he certainly didn’t consider himself stupid. He also explicitly claims that his conclusions concerning God’s natural and moral attributes are the result of an analogical design argument. Moreover, when he mentions the atheists, who certainly wouldn’t consent to having any such universal propensity to believe in God, he claims that their atheism is only nominal because even they would admit there to be “a certain degree of analogy among all the operations of Nature.” (DNR 12.8; KS 218) Again, he is mentioning the analogical argument for design, but we don’t see

17 Very interesting ideas on these topics could be found in Gaskin 1988, Part One, Ch. 7; McCormick 1993; Lorkowski 2014.

him returning to the irregular argument from Part 3, which is why we have to look for a different explanation of his alleged reversal.

After Philo's initial confession a large portion of Part 12 of the *Dialogues* is dedicated to discussing the question of practical consequences of religion. Along with his "veneration for true religion" Philo also expresses his "abhorrence of vulgar superstitions", to which Cleanthes retorts that "religion, however corrupted, is still better than no religion at all." (DNR 12.9–12.10; KS 219–20) Philo strongly disagrees with this and claims that religion "as it has commonly been found in the world" has a rather bad influence on human society and its morality, often bringing about creation of various factions, persecutions, civil wars, oppression and generally nurturing religious fanatics who act immorally in order to achieve some higher religiously inspired end (DNR 12.11; KS 220). Upon hearing this and realising that Philo's stance towards religion is still highly critical, Cleanthes warns him: "take care: push not matters too far: allow not your zeal against false religion to undermine your veneration for the true." (DNR 12.24; KS 224) However, it soon becomes clear that Philo and Cleanthes don't have the same idea in mind when they speak of "true religion", for Cleanthes then expresses the following conviction:

"The most agreeable reflection, which it is possible for human imagination to suggest, is that of genuine Theism, which represents us as the workmanship of a Being perfectly good, wise, and powerful; who created us for happiness, and who, having implanted in us immeasurable desires of good, will prolong our existence to all eternity, and will transfer us into an infinite variety of scenes, in order to satisfy those desires, and render our felicity complete and durable." (*ibid.*)

To the careful reader of the *Dialogues* this should sound almost as surprising as "Philo's reversal" because "[t]he contrast of Cleanthes' expression of adherence to some form of 'genuine theism' with his earlier accounts of the argument from design is remarkable." (Lemmens 2012a: 290) Lemmens thinks that, although his defence of this argument seems at first to be separated from traditional Christianity, Cleanthes in Part 12 "links his experimental theology explicitly with the apologetic purposes he had in mind the whole time." (Lemmens 2012a, 285) I am not sure whether this is the best way to explain Cleanthes' reversal because to me it seems that he actually dismisses experimental theology rather than linking it to his Christian beliefs. In this part of the *Dialogues* he even claims that "[t]he doctrine of a future state is so strong and necessary a security to morals, that we never ought to abandon or neglect it" (DNR 12.10; KS 220) and so it becomes evident that his God is the traditional Christian God who cares about our well-being. Cleanthes now describes him as "perfectly good, wise, and powerful" although he previously conceded that God could only be finitely perfect. Because this form of theism cannot be derived from the argument from design, for nothing in our experience of the world warrants the beliefs in personal providential deity

of traditional Christianity, I believe that Lemmens is right to conclude that “there is a conceptual gap lurking between the argument from design [...] and the genuine providential theism unfolded in the last part.” (Lemmens 2012a, 291) This discovery makes the fact that Cleanthes was unusually unperturbed by Philo’s rather successful attacks on his argument from design less surprising, because it now becomes clear that his faith never rested on any experiential grounds.

Philo, on the other hand, rejects the idea that eternal reward and punishment are the source of human morality. Instead he says that “the smallest grain of natural honesty and benevolence has more effect on men’s conduct, than the most pompous views, suggested by theological theories and systems” (DNR 12.13; KS 221), thus pointing towards Hume’s well-known moral theory according to which the source of human morality lies in human nature and the sentiments of approval and disapproval rather than religion. He agrees with Cleanthes’ claim that religion offers us consolation, but he explains this to be the case because “terror is the primary principle of religion” and he further claims that “both fear and hope enter into religion” (DNR 12.29; KS 225–226), which is exactly what Hume says in *The Natural History of Religion* concerning the origin of religion in human nature (NHR 2.4). When we look closer, Part 12 of the *Dialogues* abounds with evidence that Philo is actually defending Hume’s views on religion, while Cleanthes’ reversal<sup>18</sup> shows that the same could not be said of him. I now turn to the conclusion of this part in order to provide the last piece of evidence for the claim that Philo remains consistent throughout the whole *Dialogues*.

### Philo’s final speech in the *Dialogues*

In his last speech in the *Dialogues* Philo says:

“If the whole of Natural Theology, as some people seem to maintain, resolves itself into one simple, though somewhat ambiguous, at least undefined proposition, *That the cause or causes of order in the universe probably bear some remote analogy to human intelligence*: If this proposition be not capable of extension, variation, or more particular explication: If it affords no inference that affects human life, or can be the source of any action or forbearance: And if the analogy, imperfect as it is, can be carried no farther than to the human intelligence; and cannot be transferred, with any appearance of probability, to the other qualities of the mind: If this really be the case, what can the

---

18 Surprisingly, Cleanthes’ reversal was almost entirely overlooked in the secondary literature on the *Dialogues* until Lemmens drew attention to it (in Lemmens 2012a, 2012b).

most inquisitive, contemplative, and religious man do more than give a plain, philosophical assent to the proposition, as often as it occurs; and believe that the arguments, on which it is established, exceed the objections, which lie against it?" (DNR 12.33; KS 227)

It looks like the most Philo is willing to accept is only a minimal conclusion of the design argument – the claim that there is *probably* some *remote* analogy between the cause or causes of order in the universe and human intelligence. The words “probably” and “remote” suggest that he accepts a very limited claim, which he also characterises as “ambiguous” and “undefined” and he limits it further by saying that he accepts it only insofar as it doesn’t affect human life. Philo’s God is obviously not the personal God of Christianity, he only stands for a vague idea of some cause or causes of the world which is/are in some ways similar to human mind and completely indifferent when it comes to human morality and our practical lives.

Moreover, we shouldn’t attach much importance to the fact that Philo accepts that there is some analogy or similarity between this “cause or causes” and human intelligence. He already said before that there is “a certain degree of analogy among all the operations of Nature” so that “the rotting of a turnip, the generation of an animal, and the structure of human thought” may be some “energies that probably bear some remote analogy to each other.” (DNR 12.8; KS 218) He here repeats the words “probably” and “remote” and his irony couldn’t be more obvious.<sup>19</sup> He likewise suggests that the whole dispute concerning theism should be understood as merely verbal. The theist would certainly agree that the analogy between human and divine intelligence is very remote, since God’s intelligence must be infinitely superior to human, while the atheist would consent to there being some kind of analogy between human and divine mind, because there is a certain degree of analogy among everything that exists. Their disagreement, therefore, concerns only the “degrees of quality” that cannot be exactly measured, which is why this dispute can never reach a final conclusion. (DNR 12.7; KS 217–219) I think that Penelhum is right in saying that “the message throughout these paragraphs is that the analogies on which the Design argument hinges are too vague to be worth denying, and therefore are too inconsequential to constitute significant assertions.” (Penelhum 2012, 208) The consequences of this explanation with regard to Philo’s “true religion” are clear. If there is some remote analogy between everything in nature, then the claim that there is this unspecified analogy between human and divine mind doesn’t say anything meaningful, and if “true religion”

---

19 If we accept the interpretation of Philo’s reversal and Philo’s final conclusion in DNR 12 which I have offered in this paper, then we also ought to interpret Philo’s recommendation of sceptical fideism at the end of the *Dialogues* as deeply ironical, because he obviously doesn’t consider classical Christianity to be a very desirable or recommendable religious standpoint.

comes down to the acceptance of one such conclusion, then this religion “is not a very *religious* religion.” (Penelhum 2012, 209)

## Conclusion

Finally, I would like to briefly touch on the question that is not the subject of my paper but that is very closely connected to it. If the meaning of Philo’s true religion is really what I attempted to establish above, and if he represents Hume in the *Dialogues*, how should we best characterise Hume’s views on religion? Many among Hume’s contemporaries considered him to be an atheist, and many interpreters today consider him to be an undercover or closeted atheist because he doesn’t openly admit his atheism (Cordry 2011, Penelhum 2008, 2012, Lemmens 2012). However, a careful reading of his writings shows that he doesn’t accept it even implicitly. He attacks different arguments for theism but he almost never tries to prove that their conclusions are wrong. He mainly insists on our having insufficient evidence to claim that they are true. On the other hand, and mostly because of Philo’s final speech Hume was characterised as a deist, and Gaskin famously described his deism as *attenuated* (Gaskin 1988). While he did in the end accept that the cause or causes of order in the universe bear some remote analogy to human mind, this conclusion is so ambiguous and weak that I don’t think it gives us good reasons to call him a deist. The way he characterised them, there might be a plurality of such causes and they might not be any kind of deity after all. We could describe Hume’s position as irreligion (Russell 2016), because Philo deprives religion of any truly religious content, but I think it would be best to describe it in the same way he himself describes his position in other places – as scepticism. Hume believes that our enquiries often take us too far, especially when we concern ourselves with religion, and then we simply don’t know what to say:

“When we carry our speculations into the two eternities, before and after the present state of things; into the creation and formation of the universe; the existence and properties of spirits; the powers and operations of one universal spirit, existing without beginning and without end; omnipotent, omniscient, immutable, infinite, and incomprehensible: We must be far removed from the smallest tendency to scepticism not to be apprehensive, that we have here got quite beyond the reach of our faculties.” (DNR 1.10; KS 134)

In claiming that one very ambiguous proposition is the whole content of true religion, Hume is telling us that there seems to be nothing more determinate we could say about that subject.

## References to Hume's works

- T Hume, David. 2009. *A Treatise of Human Nature*. Edited by David Fate Norton and Mary J. Norton. Oxford: Oxford University Press. References are to to book, part, section, and paragraph, followed by page references to the Selby-Bigge/Nidditch edition.
- Hume, David. 1978. *A Treatise of Human Nature*. Edited by L. A. Selby-Bigge, rev. by P. H. Nidditch. Oxford: Clarendon Press.
- EHU Hume, David. 1975. *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. Edited by L. A. Selby-Bigge, rev. by P. H. Nidditch. Oxford: Clarendon Press.
- Hume, David. 2000. *An Enquiry concerning Human Understanding*. Edited by Tom L. Beauchamp. Oxford: Clarendon Press. References are to section and paragraph, followed by page references to the Selby-Bigge/Nidditch edition.
- DNR Hume, David. 1948. *Dialogues concerning Natural Religion*. Edited by Norman Kemp Smith. 2nd ed. New York: Social Science Publishers. References are to part and paragraph, followed by page references.
- NHR Hume, David. 2007. *The Natural History of Religion*. In *A Dissertation on the Passions and the Natural History of Religion*. Edited by Tom L. Beauchamp. Oxford: Oxford University Press. References are to section and paragraph.
- LG Hume, David. 2007. *A Letter from a Gentleman to his Friend in Edinburgh*. In (prior.) *An Enquiry Concerning Human Understanding and Other Writings*, Edited by Stephen Buckle, 147–162. Cambridge: Cambridge University Press. References are to page number.
- LH Greig, J. Y. T. 1932. *The Letters of David Hume*, 2 Vols. Oxford: Clarendon Press. References are to volume and page number.

## References to other works

- Battersby, Christine. 1979. "The *Dialogues* as original imitation: Cicero and the nature of Hume's scepticism." In *McGill Hume Studies*, edited by David F. Norton, Nicholas Capaldi and Wade L. Robison, 239–52. San Diego: Austin Hill Press.
- Bricke, John. 1975. "On the Interpretation of Hume's *Dialogues*." *Religious Studies* 11 (1): 1–18.
- Cicero, Marcus Tullius. 45 BCE/1997. *The Nature of the Gods*. Translated by PG. Walsh. Oxford: Oxford University Press.

- Cordry, Benjamin S. 2011. "A more dangerous enemy? Philo's 'confession' and Hume's soft atheism." *International Journal for Philosophy of Religion* 70: 61–83.
- Dye, James. 1989. "A Word on Behalf of Demea." *Hume Studies* 15: 120–40.
- Dye, James. 1992. "Demea's Departure." *Hume Studies* 18: 467–81.
- Fogelin, Robert J. 1983. "The Tendency of Hume's Skepticism." In *The Skeptical Tradition*, edited by Myles Burnyeat, 397–412. Berkley: University of California Press.
- Fogelin, Robert. 2017. *Hume's presence in the Dialogues concerning natural religion*. New York: Oxford University Press.
- Foley, Rich. 2006. "Unnatural Religion: Indoctrination and Philo's Reversal in Hume's *Dialogues* Concerning Natural Religion." *Hume Studies* 32 (1): 83–112.
- Fosl, Peter S. 1994. "Doubt and Divinity: Cicero's Influence on Hume's Religious Skepticism." *Hume Studies* 20 (1): 103–120.
- Gaskin, J. C. A. 1988. *Hume's Philosophy of Religion*. Second Edition, London: Palgrave-MacMillan.
- Harward, D. W. 1975. "Hume's *Dialogues* Revisited." *International Journal for Philosophy of Religion* 6: 137–153.
- Kemp Smith, "Introduction", In *Dialogues concerning Natural Religion*. Edited by Norman Kemp Smith, 1–75. 2nd ed. New York: Social Science Publishers.
- Lemmens, Willem. 2012a. "Hume's Atheistic Agenda: Philo's Confession in *Dialogues* 12." *Bijdragen, International Journal in Philosophy and Theology* 73 (3): 281–303.
- Lemmens, Willem. 2012b. "The 'true religion' of the sceptic: Penelhum reading Hume's *Dialogues*." *Canadian Journal of Philosophy* 42 (S1): 183–197.
- Lorkowski, Chris M. 2014. "Attenuated Deism and Hume's Propensity to Believe." *Society and Politics* 8 (1): 60–77.
- Lorkowski, Chris M. 2016. "Doxastic Naturalism and Hume's Voice in the *Dialogues*." *The Journal of Scottish Philosophy* 14 (3): 253–274.
- McCormick, Miriam. 1993. "Hume on Natural Belief and Original Principles." *Hume Studies* 29 (1): 103–116.
- Mossner, Ernest C. 1936. "The Enigma of Hume." *Mind*, NS 45 (179): 334–349.
- Olszewsky, Thomas M. 2003. "Demea's Dilemmas." *British Journal for the History of Philosophy* 11(3): 473–492.



- Penelhum, Terence. 2012. "Hume's atheism and the role of Cleanthes." *Canadian Journal of Philosophy* 42 (S1): 206–211.
- Price, John Valdimir. 1964. "Sceptics in Cicero and Hume." *Journal of the History of Ideas*, Vol. 25, No. 1: 97–106.
- Rasmussen, Dennis C. 2017 *The Infidel and the Professor: David Hume, Adam Smith and the Friendship that Shaped Modern Thought*. New Jersey: Princeton University Press.
- Russell, Paul. 2016. "Hume's Philosophy of Irreligion and the Myth of British Empiricism." In *The Oxford Handbook of David Hume*, edited by Paul Russell, 109–137. New York: Oxford University Press.
- Stahl, Donald. 1984. "Hume's Dialogue IX defended." *Philosophical Quarterly* 34: 505–507.
- Stewart, M. A. 2000. "The Dating of Hume's Principal Manuscripts." In *The Scottish Enlightenment: Essays in Reinterpretation*, Edited by Paul Wood, 267–314. Rochester NY: University of Rochester Press.
- Stove, D. C. 1978. "Part IX of Hume's *Dialogues*." *Philosophical Quarterly* 28: 300–309.
- Vink, A. G. 1986. "The Literary and Dramatic Character of Hume's *Dialogues Concerning Natural Religion*." *Religious Studies* 22 (3–4): 387–396.
- Wieand, Jeffery. 1985. "Pamphilus in Hume's *Dialogues*." *The Journal of Religion* 65 (1): 33–45.



## THE ROLE OF THE NATURAL HISTORY OF RELIGION IN HUME'S CRITIQUE OF RELIGIOUS BELIEF

**Abstract:** *I argue that Hume's naturalistic explanation of religious belief in the Natural History of Religion has significant epistemic consequences. While he argues in the Dialogues Concerning Natural Religion (and in other works) that belief in God is not justified on the basis of testimony or philosophical argument, this is not enough to show that religious belief is not warranted. In the Natural History, Hume provides a genetic explanation for religious belief. I contend that the explanation of religious belief in the Natural History, given Hume's conclusions in his other works, provides grounds to reject religious belief. Thus, I conclude that the Natural History plays an important role in Hume's overall critique of religion insofar as it is a necessary component in Hume's arsenal against the warrant of religious belief.*

**Keywords:** *Hume, Natural History of Religion, Religion, Religious Belief, Dialogues Concerning Natural Religion, Natural Belief, Treatise Of Human Nature, Warrant, Propensity, Anthropomorphize, Adulate, Monotheism, Polytheism*

### 1. Introduction<sup>1</sup>

Hume opens the *Natural History of Religion*:

As every enquiry, which regards religion, is of the utmost importance, there are two questions in particular, which challenge our attention, to wit, that concerning its foundation in reason, and that concerning its origin in human nature (NHR Introduction.1; 134).

Hume's works on religion—the *Natural History* and the *Dialogues*—are often interpreted as two separate projects—the *Dialogues* an investigation

---

1 References to *A Treatise of Human Nature* are to the Clarendon edition (Hume 2000b), abbreviated *T* in the text and cited by book, part, section, paragraph number, followed by the page number from the Selby-Bigge/Nidditch edition (Hume 1978). For example (*T* 2.3.9.3; 438). References to *Dialogues Concerning Natural Religion* are to the Gaskin edition (Hume 1993), abbreviated *DNR* in the text and cited by the section number, followed by the page number from the Kemp Smith edition (Hume 1947). For example (*DNR* X; 95; 193). References to the *Natural History of Religion* are from the Clarendon edition (Hume 2007), abbreviated *NHR* in the text and cited by section and paragraph number, followed by the page number from the Gaskin edition (Hume 1993). For example (*NHR* 1.3; 135). References to the *Enquiry Concerning Human Understanding* are to the Clarendon edition (Hume 2000a) and cited by section, part, and paragraph number. For example, (EHU 4.1.5).

of religion's foundation in reason; the *Natural History* an investigation of religion's origin in human nature. However, I argue that these two works contribute to a unified epistemological project. Hume intends the explanation he provides in the *Natural History* to have negative epistemic consequences insofar as it shows that religious belief—be it polytheistic or monotheistic—is not a “natural belief”.<sup>2</sup> Taken in conjunction with his arguments in *Dialogues* and in some of his other texts, it justifies rejecting religious belief. Thus, the *Natural History* plays an indispensable role in Hume's critique of the epistemic status of religious belief.

In Section 2, I briefly recap the conclusions of Hume's other works on religion, including the essay “Of Miracles” and the *Dialogues*. For the purposes of this paper, I assume that the result of these taken together is that, according to Hume, religious belief is not justified by testimony or cogent philosophical argument.<sup>3</sup> In Section 3, I note that given a naturalistic interpretation of Hume's epistemology, the fact that religious belief is not justified on the basis of philosophical argument is not enough to show that religious belief should be rejected—it may be a natural belief and therefore enjoy positive epistemic status. In order to determine whether religious belief is a natural belief, we must investigate its genesis, determining whether the propensities responsible

- 
- 2 By arguing that Hume has an epistemic goal in the *Natural History*, I do not intend to deny that Hume has other goals. For example, Mark Webb argues that Hume's goal in the *Natural History* is to convince his audience to abandon traditional Christian practice; in particular, he contends that the arguments in the *Natural History* purport to show the moral inferiority of popular religion, in part by showing that polytheism is morally preferable to monotheism (Webb (1991) 150ff). Webb's account not only coheres with what Hume has to say in his other discussions of religion—in particular, with several of Philo's comments in *Dialogues* part XII—but also presents an excellent interpretation of parts IX through XV of the *Natural History*. Critiquing the effect on morality caused by popular religion was certainly important to Hume. So, while I agree that the moral critique plays an important role in Hume's goals in the *Natural History*, I disagree with Webb's claim that it “is *principally* a moral critique aimed at severing allegiance to traditional religion” (Webb (1991) 149, my emphasis). Webb's interpretation does not do justice to Hume's aims in sections I-VIII of the *Natural History*. He claims that Hume's account of the causes of religious belief is merely the means to an end—the end of severing allegiance to traditional religious practice: “while the work purports to be a natural history of religious beliefs and practices, suggesting something of a ‘scientific’ examination of a ‘natural’ phenomenon, that feature of the work clearly is intended by Hume to facilitate his moral critique of traditional religious belief” (Webb (1991) 145). But I will argue that this is not the case: Hume's explanation of the origin of religious belief has important epistemic consequences. While I do agree that Hume's moral critique of religious belief is an important aspect of the *Natural History*, Hume's account of the causal origin of religious belief is not merely in the service of this goal.
- 3 Hume's focus in these texts is centered on refuting arguments for monotheism (and Christianity in particular), as he mostly takes for granted the falsity of polytheism. However, much of what he says would be applicable to arguments or testimony in favor of robust polytheism. I discuss Hume's attitude toward polytheism in more detail in my “Hume on the Epistemic Superiority of Polytheism to Monotheism” (unpublished manuscript).

for the belief are similar to custom or more like those that Hume contrasts negatively with custom. In Section 4, I examine the primary psychological principles responsible for religious belief and argue that Hume's explanation of religious belief has negative epistemic consequences. Not only is religious belief not natural, but given its causes, we have reason to reject it. In Section 5, I conclude that the *Natural History* plays a crucial role in Hume's overall critique of religious belief.

## 2. Hume's Conclusions in *The Dialogues* and Other Works

The *Dialogues Concerning Natural Religion*, along with other relevant texts, such as the essay "Of Miracles" (EHU 10), provide a scathing critique of the justification of religious belief. While a full survey of Hume's other works is beyond the scope of this paper, I will summarize the highlights below.

In the essay "Of Miracles", Hume argues that religious belief cannot be justified by testimonial evidence. He concludes, "we may establish as a maxim, that no human testimony can have such force as to prove a miracle, and make it a just foundation for any such system of religion" (EHU 10.2.35).

In the essay "Of a Particular Providence and of a Future State", Hume denies the justification for belief in "a providence...and supreme governor of the world, who guides the course of events, and punishes the virtuous with honor and success..." (EHU 11.20) on the grounds that this inference is not warranted on the basis of an inference from known effects to causes, but rather a fallacious inference ("a gross sophism" (EHU 11.20)) from an assumed cause (a benevolent deity) to expected effects.

In the *Enquiry* as well as in the *Treatise*, Hume rejects all *a priori* arguments in support of "matters of fact"—truths the negations of which are conceivable.<sup>4</sup> For Hume, *a priori* arguments for God's existence are rejected because the "matter of fact" (that God exists) is not demonstrable.<sup>5</sup> Hume reiterates this position in the *Dialogues*:

I shall begin with observing, that there is an evident absurdity in pretending to demonstrate a "matter of fact", or to prove it by any arguments *a priori*. Nothing is demonstrable, unless the contrary implies a contradiction. Nothing, that is distinctly conceivable, implies a contradiction. Whatever we conceive as existent, we can also conceive as non-existent. There is no Being, therefore, whose non-existence implies a contradiction. Consequently there is no Being, whose existence is demonstrable. I propose this argument as entirely decisive, and am willing to rest the whole controversy upon it (DNR IX; 91; 189).

---

4 See EHU 4.1.6ff. and also T 1.3.1.1ff.; 69ff.

5 Moreover, in the *Treatise*, Hume implicitly rejects the traditional ontological argument insofar as he argues existence is not a property or predicate (T 1.2.6.3–5; 66–67).

Hume devotes an entire work—the *Dialogues*—to the argument from design—one of the two traditional *a posteriori* arguments for God’s existence.<sup>6</sup> The *Dialogues* provides numerous forceful objections against the *a posteriori* argument from design: e.g., that the analogy between the universe and a machine is weak (DNR II); that the argument (at best) allows us to infer an imperfect deity (DNR V), and would also justify us in inferring a multitude of deities (DNR V); that the analogical argument does not establish that the principle governing order is intelligence, rather than vegetation or generation (DNR VII); and that there are alternative (Epicurean and proto-Darwinian) explanations of the order found in the universe (DNR VIII). The *Dialogues*, as a whole, conclude that belief in God cannot be justified by the *a posteriori* argument from design.<sup>7</sup>

For the purposes of this paper, I will operate on the assumption that, taken together, Hume’s arguments against religious belief found in the *Dialogues* (and other texts besides the *Natural History*) show that it cannot be justified on the basis of testimony, or on *a priori* or *a posteriori* argument.

### 3. Natural Belief and the Explanation of False, Meaningless, or Unjustified Beliefs

For Hume the naturalist, the fact that religious belief is not justified by philosophical argument is not enough to show that it should be rejected. After all, Hume showed that beliefs based on induction are not justified by argument but still have positive epistemic status. A predominant view among naturalist interpreters of Hume suggests that he holds that some beliefs—

6 Hume also argues against the cosmological argument in Part IX of the *Dialogues*.

7 The most pressing challenge to this interpretation is the problem of “Philo’s Reversal” in *Dialogues* XII. J.C.A. Gaskin, for example, argues that “he [Philo] concedes that belief in a designer is after all a rational belief” (Gaskin 1988, 138), and concludes that Hume therefore endorses “attenuated deism”: the view that the argument from design justifies the belief in a designer of the universe who possesses “the power of an agent,” and who bears “some remote analogy to human intelligence,” but “cannot be known to have any moral attributes” (Gaskin (1988) 167). Other commentators (e.g., Andre, Holley, Immerwahr) are sympathetic to this view, though some go on to argue that all things considered, Hume does not endorse theism. Still other commentators (Falkenstein, Garrett, Black and Gressis, Willis, etc.) argue that something a bit more robust—“genuine,” “true,” or “philosophical theism” is warranted, though the details vary. I argue against “genuine theism” views, offering an alternative interpretation of Philo’s Reversal in my “Philo’s Reversal and Hume on Belief” (unpublished manuscript) where I conclude that the supposed belief Philo endorses is so qualified and so void of content (meaningless)—that it does not constitute a belief according to Hume’s understanding of that term. Russell 2008 and Lemmens 2012 argue for a similar conclusion. Kraal 2013 argues for a related conclusion, claiming that *Dialogues* X shows that moral terms applied to God are meaningless. Holden 2010 also argues for a related conclusion—that in Hume’s view no first cause could possibly have moral attributes, and the existence of a first cause would not have any implications for human behavior).

*natural beliefs*— are warranted<sup>8</sup> by features of the belief-forming mechanisms that produce them, and that he would endorse a naturalistic theory of warrant—perhaps reliabilism, or a proper function theory—to sustain this position.<sup>9</sup> Here, I take it for granted that there are natural beliefs, and that beliefs based on induction are among them. Given this broad interpretive framework, religious belief, some argue, is yet another Humean natural belief—like induction, warranted in virtue of the mechanisms that cause it.<sup>10</sup>

It should be noted, however, that Hume notes that not all beliefs based on custom are natural beliefs, insofar as some lack warrant, based on certain circumstances that may occur in the belief-forming process: when the objects are not perfectly resembling or constantly conjoined, when a counterexample in experience happened in the distant past but confirming experiences happened more recently, and when we rely on general rules (T 1.3.13 1–9; 143–147). This problem, Hume claims, can be corrected by ensuring adherence to several “rules by which to judge of causes and effects” such as ensuring that the cause and effect are contiguous in space and time, with the cause being prior to the effect, only endorsing the causal claim when cause and effect are constantly conjoined, etc. (T 1.3.15; 173–176).

Like beliefs based on induction, religious belief is widespread. Hume admits that belief in “invisible, intelligent power has been very generally diffused over the human race, in all places and in all ages” (NHR Introduction.1; 134). He argues in the *Natural History* that religious belief, like inductive belief, is the result of several natural propensities of the imagination. But having natural causes does not guarantee that a belief is a natural belief (in the technical sense) since every belief has a natural cause. Just because Hume provides a naturalistic explanation of a belief does not entail that it is a natural belief.

In the *Treatise*, Hume frequently provides causal explanations of unjustified, false, or meaningless beliefs, often after pointing out or arguing for a defect in the justification or content of the belief. This is not surprising, since once he has shown that a belief is defective, certain questions arise: Why do people believe what is obviously false? If the belief is not justified

---

8 While commentators differ on how to describe beliefs that possess positive epistemic status, I will call them *warranted*, following Plantinga who introduced the term as a “name for that property—or better, *quantity*—enough of which is what makes the difference between knowledge and mere true belief” (Plantinga (2000) xi).

9 The term natural belief was coined by Kemp Smith in his “The Naturalism of Hume I and II,” (Kemp Smith 1905 a and b) and elaborated upon in his *The Philosophy of David Hume* (Kemp Smith 1941). Some other commentators who hold this view include: Arnold 1983, Beauchamp & Mappes 1975, Beauchamp & Rosenberg 1981, Broughton 1983, Connon 1976, Craig 1987, Fogelin 1985, Lenz 1958, Loeb 2002, Schmitt 1992, Spector 2003, Winters 1979, and Wolterstorff 1996.

10 See Butler 1960, Costelloe 2004, Harris 1987, Hurlbutt 1985, Logan 1992, 1993 and 1996, Noxon 1995, Penelhum 1979 (though he changes his view in Penelhum 1983), Pike 1970, Prado 1981, Reich 1998, and Tweyman 1986.

by evidence or argument, on what basis do people hold the belief? Because Hume's project is to understand human nature, he is interested in answering these questions.

For example, he explains the source of the vulgar belief in external existence, which is not justified by the senses or reason (T 1.4.2.3–14; 188–193). Belief in body is partly caused by the propensity to ascribe identity to related objects (T 1.4.2.24, 31–35; 199, 202–204). In both the *Treatise* and the first *Enquiry*, immediately following Hume's skeptical attack on the justification of inductive beliefs, he provides a naturalistic explanation for them (T 1.3.6.12–15 and EHU 5.1). They're caused by the propensity to believe in (or expect) the unobserved effect of an observed cause (or to believe in the unobserved cause of an observed effect) given previous experience of their constant conjunction.

However, not all beliefs for which Hume provides a causal explanation are natural beliefs. For example, he explains the origin of the philosophers' system of "double existence," a "monstrous offspring of two [opposing] principles," by appealing to a compulsion to "set ourselves at ease as much as possible" (T 1.4.2.52; 215). He grounds the ancients' beliefs in "sympathies, antipathies, and horrors of a vacuum"—beliefs he calls "fictions"—in the propensity to anthropomorphize (T 1.4.3.11; 224). He says that the "unreasonable and capricious" "fiction" (T 1.4.3.1; 219)—belief in material substance (T 1.4.3.2–8; 219–222)—is "derived from principles as natural as any of those above-explain'd" (T 1.4.3.2.8; 222).

Hume thinks inductive beliefs have positive epistemic status (or, as I would put it, warrant),<sup>11</sup> but some beliefs for which he gives naturalistic explanations do not. According to the naturalistic interpretation of Hume, it is because of the *naturalness* of beliefs based on induction that Hume deems them warranted: "tho' causation be a *philosophical* relation, as implying contiguity, succession, and constant conjunction, yet 'tis only so far as it is a *natural* relation, and produces an union among our ideas, that we are able to reason upon it, or draw any inference from it" (T 1.3.6.16; 94). He

11 Before the main arguments concerning induction in Book I, Part II, Section VI, Hume claims that "the relation of cause and effect" is the only relation that allows us to "go beyond what is immediately present to the senses, either to *discover* the real existence or the relations of objects" (T 1.3.2.2; 73, emphasis added). He says that the relation of causation "*informs* us of existences and objects which we do not see or feel" (T 1.3.2.3; 74, emphasis added). Within Section VI, he claims that the relation of causation "'tis the only one, on which we can found a *just* inference from one object to another" (T 1.3.6.7; 89, emphasis added). After Section VI, he not only refers to inductive inference as a "true species of reasoning" (T 1.3.7.5 n20; 97), but also states, "A person, who stops short in his journey upon meeting a river in his way, *foresees* the consequences of his proceeding forward; and his *knowledge* of these consequences is convey'd to him by past experience, which *informs* him of such certain conjunctions of causes and effects" (T 1.3.8.13; 103, emphasis added). Finally, Hume claims, "one who concludes somebody to be near him, when he hears an articulate voice in the dark, reasons *justly* and naturally" (T 1.4.4.1; 225, emphasis added).



distinguishes between natural beliefs and unwarranted beliefs on the basis of certain features of their causes—while some belief-forming mechanisms confer warrant, others do not. He explains:

I must distinguish in the imagination betwixt the principles which are permanent, irresistible, and universal; such as the customary transition from causes to effects, and from effects to causes: And the principles, which are changeable, weak, and irregular;[...] The former are the foundation of all our thoughts and actions, so that upon their removal human nature must immediately perish and go to ruin. The latter are neither unavoidable to mankind, nor necessary, or so much as useful in the conduct of life; but on the contrary are observ'd only to take place in weak minds, and being opposite to the other principles of custom and reasoning, may easily be subverted by a due contrast and opposition. For this reason the former are received by philosophy, and the latter rejected. (T 1.3.4.1; 225)

Those beliefs caused by permanent, irresistible, and universal principles are warranted natural beliefs; those beliefs caused by changeable, weak, and irregular principles are not. Hume makes this explicit in the case of the opinions of the ancient philosophers.<sup>12</sup> Thus, in order to determine whether a belief is among the natural beliefs, we must look for more than the mere existence of a causal explanation of the belief.

Because natural beliefs have positive epistemic status, one way to determine whether a belief is a natural belief is to examine directly what Hume says about the epistemic status of the belief. If he expresses epistemic approval, this is evidence that the belief in question is a natural belief; if he expresses epistemic disapproval, then this is evidence that the belief is not natural. Depending on his language, this may also be evidence that the belief should be rejected. For example, if Hume says a belief is false, then this is evidence that he believes we should avoid or reject it.

Additionally, natural beliefs are warranted by features of the psychological propensities that cause them; so, another way to determine whether a belief is natural is to examine its particular cause(s). If the propensity responsible for the belief is not irresistible, universal, or permanent, then the belief is not natural. If the propensity in question causes other beliefs that Hume impugns epistemically, then we have evidence that no belief caused by that propensity is natural. If the belief-forming mechanism is explicitly contrasted with custom, then a belief formed by it is not a natural belief. If the propensity can

12 "But one, who is tormented he knows not why, with the apprehension of specters in the dark, may, perhaps, be said to reason, and to reason naturally too: But then it must be in the same sense, that a malady is said to be natural; as arising from natural causes, tho' it be contrary to health, the most agreeable and most natural situation of man. The opinions of the antient philosophers...are like the specters in the dark, and are deriv'd from principles, which, however common, are neither universal nor unavoidable in human nature" (T 1.4.4.1–2; 225–226).

or should be resisted or subverted, or if it is weak or irregular, beliefs caused by it should be withheld or rejected.

To determine the epistemic status of religious belief, then we must turn to the *Natural History* where Hume's stated goal is to determine religion's "origin in human nature" (NHR Introduction.1; 134). Hume's project in the *Natural History* mirrors his treatment of the beliefs of the ancient philosophers, but also beliefs based on induction. This provides an opening for the view that Hume thinks religious beliefs are warranted. On this view, Hume's naturalistic explanation of religious belief in the *Natural History* has *positive* epistemic consequences. I argue that this is not the case. I show that the propensities responsible for religious belief do not generate natural beliefs and that the language Hume uses demonstrates his epistemic disapproval. I argue that Hume's naturalistic explanation of religious belief has *negative* epistemic consequences and conclude that this gives us reason to reject religious belief.

#### 4. The Explanation of Religious Belief in the *Natural History*

In the *Natural History*, Hume provides a causal account of religious belief. He claims early religious belief is polytheistic and over many generations is transformed into monotheism. He asserts that religious belief is caused by several propensities—including the propensity "toward a system that provides satisfaction" (NHR 3.1–2; 141), the propensity to anthropomorphize and the propensity to adulate. Hume claims both polytheism and monotheism originate from the belief in invisible, intelligent power ("the original belief"),<sup>13</sup> caused by the propensity toward a system triggered by the disorder in nature. Other propensities, including anthropomorphizing and adulation, operate on the belief in invisible intelligent power, causing first robust polytheism and eventually monotheism. Throughout, Hume emphasizes the role of fear and ignorance, along with the desire to relieve related anxieties. In what follows, I show that the propensities active in Hume's account have negative epistemic consequences for the beliefs they generate.<sup>14</sup>

13 Kail refers to the belief in invisible intelligent power as the "core concept," and notes that all religious belief, according to Hume, stems from this original belief (Kail 2000a 192 and 2000b 7).

14 Kail (2007b) has a different, but I think complimentary, interpretation of the *Natural History*. He claims that Hume's explanation in the NHR provides reason to suspend all religious belief insofar as it provides reason to suspend belief in invisible, intelligent power (the core concept), the content of which is necessary and sufficient for any religious belief. He claims that Hume's line of reasoning is aimed at the fideist (who believes that one can be rational in maintaining religious belief even if there are no arguments, reasons, or evidence in favor of it) insofar as it provides reason to suspend belief without attacking arguments or evidence in its favor. I will claim, instead, that it provides reason to reject and not just suspend religious belief. For someone (like the fideist) who has not already examined the arguments for and against religious belief, they would need to do so before rejecting that belief, as Kail suggests. Hume, on the other hand, typically only

#### 4a. *The Propensity Toward a System that Provides Satisfaction*

Hume describes the propensity responsible for the first occurrence of belief in invisible, intelligent powers as “a propensity in human nature, which leads to a system, that gives them some satisfaction” (NHR 3.1–2; 141). This propensity functions as a means of coping with the situation in which man<sup>15</sup> finds himself—terrified and dependent on the world around him—by leading him to feel that he has some control over his own destiny. He fears the unknown causes of calamitous natural events; he hopes that the course of nature will unfold in such a way to keep him safe, secure, and well-fed. By positing that nature is controlled by intelligent powers, man can more easily believe that he has a means of controlling his fate. For example, he can provide sacrifices in order to please the intelligent powers. This reduces the anxiety produced by lack of control.<sup>16</sup> In this way, the propensity toward a system is akin to wish-fulfillment, a mechanism that produces beliefs that produce happiness or comfort.<sup>17</sup>

Both the propensity toward a system and wish-fulfillment operate in conditions where the believer is afraid, and both relieve anxiety. In both cases, an explanation is believed, but it doesn't matter how well the explanation fits the relevant data; instead, the aim of belief formation is happiness or the relief of anxiety. Like wish-fulfillment, the propensity toward a system that provides satisfaction is not truth-oriented.

From the fact that the propensity toward a system is not truth-oriented, we cannot immediately conclude that it is a propensity that does not generate natural beliefs. This fact alone does not show that it is changeable, weak, irregular, avoidable, or opposed to custom. However, if we discover that a belief is caused by a propensity that is not truth-oriented, this gives us a reason to suspend belief until we examine the evidence. Suppose, for example, I seem to remember eating potato casserole for dinner last week. I think about it for a minute, and realize that I actually dreamt that I ate potato casserole for dinner. This doesn't necessarily give me reason to reject that belief, though it does give

---

provides genetic explanations for beliefs that cannot be justified by reason, argument, or evidence. While Kail thinks Hume has a very targeted audience for the NHR, I believe the audience is broader—directed at all believers; thus anyone who has considered the arguments for and against religious belief and is searching for an alternative explanation (as Hume searches for an explanation of inductive beliefs), would see that the natural explanation for religious belief does not provide the warrant that the parallel explanation does in the case of causal beliefs.

15 In the *Natural History* and his other writings, Hume uses gendered language throughout (e.g., “man”, “mankind”). In what follows, I will often mimic Hume's gendered language; though I will sometimes use more gender-neutral terms such as “people,” proceeding on the assumption that Hume means these gendered terms to refer to all humans.

16 Jones (1972) 322, and Ferreira (1995) 595, argue for a similar interpretation.

17 Similarly, Kail (2000a 199–200 and 2000b 10–11) argues that belief in the core concept arises due to a kind of motivated irrationality.

me reason to suspend it until I verify my memory. Given that the belief in invisible intelligent power is caused by a non-truth-oriented propensity, beliefs caused by it are not natural beliefs insofar as the propensity toward a system doesn't generate beliefs with *positive* epistemic status.

Hume goes on to say that the propensity toward a system does not generate true belief. Instead, he claims that early man is *ignorant* of the true causes of natural occurrences (NHR 3.1; 141, emphasis added). Belief in invisible intelligent power as the cause of natural events is belief in something false. This evidence suggests that instead of merely suspending belief in invisible intelligent power, we should reject it as we should reject any false belief.

According to Hume, the first instantiation of the original belief is a kind of minimal polytheism (insofar as anthropomorphism has not yet begun to operate and attribute further human characteristics to the invisible intelligent powers<sup>18</sup>) (NHR 1; 135–138). Hume condemns minimal polytheism: calling it “the grossest theory” (NHR 4.12; 149) and “superstitious” (e.g., NHR 3.3, 4.2, 4.5). Since he often makes negative comments about the polytheistic system, it is clear that Hume thinks minimal polytheism, and therefore the original belief in invisible intelligent power, unwarranted.

#### 4b. *The Propensity to Anthropomorphize*

The second mechanism causally responsible for religious belief is the propensity to anthropomorphize: a “universal tendency among mankind to conceive all beings like themselves” (NHR 3.2; 141). This propensity is responsible for man's ascribing intelligence and other human characteristics to the invisible powers of minimal polytheism. While Hume says that the propensity to anthropomorphize is universal, this does not mean he thinks it cannot be resisted. He writes: “we find human faces in the moon, armies in the clouds, and by a natural propensity, *if not corrected by experience and reflection*, ascribe malice or good-will to every thing, that hurts or pleases us” (NHR 3.1; 141, emphasis added). It is clear that Hume thinks that the tendency to personify *can* and *should* be avoided, especially since it results in false beliefs, noting that in poetry, “mountains and streams are personified, and the inanimate parts of nature acquire sentiment and passion” (NHR 3.2; 141). Although this tendency is sometimes avoided, Hume makes a point to emphasize that it is not done easily: “philosophers cannot entirely exempt themselves from this natural frailty; but have oft ascribed it to inanimate matter the horror of a *vacuum*, sympathies, antipathies, and other affections of human nature” (NHR 3.2; 141).

Hume's language echoes that of *Treatise* 1.4.3, “Of the antient philosophy,” where he discusses the propensity to anthropomorphize with an even more disapproving tone. In explaining the ancient philosophers' ideas of “sympathies, antipathies, and horrors of a vacuum,” he claims, “there is a very remarkable inclination in human nature to bestow on external objects

---

18 I discuss this further in section 3b.

the same emotions, which it observes in itself; and to find every where those ideas, which are most present to it" (T 1.4.3.11; 224). Here though, Hume claims that it can be *easily* avoided, and makes clear that beliefs caused by it are unwarranted: "this inclination, 'tis true, is suppress'd by a little reflection, and only takes place in children, poets and the antient philosophers" (T 1.4.3.11; 224). Although he thinks it understandable that children personify the stones that hurt them--and poets, everything--he is not so ready to free the philosophers from blame. He asks, "we must pardon children, because of their age; poets, because they profess to follow implicitly the suggestions of their fancy: But what excuse shall we find to justify our philosophers in so signal a weakness?" (T 1.4.3.11, 225).

The paradigm example of a Humean natural belief is belief based on induction, but in the *Treatise* Hume takes care to distinguish beliefs caused by the personifying propensity from beliefs caused by custom (T 1.4.4.1; 225). This is where he differentiates between principles of the imagination that are "permanent, irresistible, and universal; *such as the customary transition from causes to effects*" and those that are "changeable, weak, and irregular; such as those I have just now taken notice of" (T 1.4.4.1; 225, emphasis added).<sup>19</sup> He refers to these principles as a defect, and contrasts them with those that "arise only from the solid, permanent, and consistent principles of the imagination" (T 1.4.4.2; 226). Hume explicitly contrasts beliefs caused by the propensity to anthropomorphize, negatively, with those caused by the inductive propensity, making clear that beliefs based on the personifying propensity are those that ought to be rejected.

Moreover, Hume says that beliefs produced by the propensity to anthropomorphize are absurd. In the *Treatise*, Hume calls beliefs based on the propensity "fictions" (T 1.4.3.11; 224). Likewise, this is the case when this propensity causes religious belief: "The absurdity is not less, while we cast our eyes upwards; and transferring, as is too usual, human passions and infirmities to the deity" (NHR 3.2; 141). As the *Treatise* predicts, in the *Natural History*, beliefs based on the personifying principle, including religious beliefs, are not warranted.

#### 4c. *The Propensity to Adulate*

Both the propensity toward a system and the propensity to anthropomorphize are responsible for polytheistic belief in invisible intelligent power. Monotheism is caused (in part) by the adulation propensity. After some time, the polytheist chooses a favorite god. The propensity to adulate causes the polytheist to elevate the favorite god, adding numerous positive qualities to his description (NHR 13.3; 176–177) until he reaches the idea of a perfect God. This tendency to raise the favorite god higher and higher is triggered by the desire to persuade the god to help him avoid misfortune and gain success (NHR 6.5; 155).

<sup>19</sup> In the first paragraph of 1.4.4, he has in view the propensity to anthropomorphize, the topics of the final paragraph of 1.4.3.

Despite the fact that men originally fear the gods and depict them as cruel, to gain the favor of their favorite god, they exalt him above the rest:

Here therefore is a kind of contradiction between the different principles of human nature, which enter into religion. Our natural terrors present the notion of a devilish and malicious deity: Our propensity to adulate leads us to acknowledge an excellent and divine. And the influence of these opposite principles are various, according to the different situation of the human understanding (NHR 13.3; 176–177).

Man's desire to relieve anxiety influences the operation of the adulation propensity. In this way, it is similar to the propensity toward a system—by raising one god to infinity, man hopes to better control his destiny: “the same anxious concern for happiness, which begets the idea of these invisible intelligent powers, allows not mankind to remain long in the first simple conception of them; as powerful, but limited beings; masters of human fate, but slaves to destiny and the course of nature” (NHR 8.2; 159). The fact that the propensity operates to relieve anxiety is evidence that it is not directed at producing true beliefs.

Moreover, Hume notes that this propensity often causes contradictions within religion: the Catholic's worship of Mary as a person equal to God, Homer's occasional tendency to honor Oceanus and Tethys as the original parents of things, and at other times bestow that honor on Jupiter, etc. (NHR 6.7–11; 156–157). He laments, “rather than relinquish this propensity to adulation, religionists, in all ages, have involved themselves in the greatest absurdities and contradictions” (NHR 6.10; 156). This indicates that at least some of the beliefs caused by the propensity are false.

But there is additional reason to reject beliefs based on the adulation propensity: Hume claims, “Men's *exaggerated praises and compliments* will swell their idea upon them; and elevating their deities to the utmost bounds of perfection, at last beget the attributes of unity and infinity, simplicity and spirituality” (NHR 8.2; 159). It is clear that, according to Hume, beliefs based on adulation are not true; instead, they are exaggerated. He also claims that the propensity to adulate leads man to attribute *false* praise to the favorite god. Because he is afraid, man must approve of what normally would be considered negative behavior or characteristics. Hume explains:

They must then be careful not to form expressly any sentiment of blame and disapprobation. All must be applause, rapture, extacy. And while their gloomy apprehensions make them ascribe to him measure of conduct, which, in human creatures, would be highly blamed, they must still affect to praise and admire that conduct in the object of their devotional addresses. (NHR 13.6; 178)

Thus, it is clear that beliefs about the favorite god, caused by the propensity to adulate, are not natural beliefs, and are false and should be avoided.

#### *4d. Monotheism Causally Dependent on Polytheism*

It is important for Hume that monotheism is not a separate belief-system, unconnected to polytheism. Rather, it is the culmination of a process of belief-transformation, beginning with polytheism and ending finally in monotheism. Over time, man is led from polytheistic belief to monotheistic belief by various psychological mechanisms. I will argue that monotheism thus inherits the epistemic problems of polytheism.

Hume's discussion of the philosopher's belief in double existence provides evidence that he thinks that epistemic problems transfer from the first belief to the second when the second belief is causally dependent on the first. According to Hume, "there are no principles either of the understanding or fancy, which lead us directly to embrace this opinion of the double existence of perceptions and objects, nor can we arrive at it but by passing thro' the common hypothesis of the identity and continuance of our interrupted perceptions [the vulgar hypothesis]" (T i.4.2.46; 211). In other words, belief in double existence is causally dependent on the vulgar belief in body. He goes on to explain, "however philosophical this new system may be esteem'd, I assert that 'tis only a palliative remedy, and that it contains all the difficulties of the vulgar system, with some others, that are peculiar to itself" (T 1.4.2.46; 211). Given that belief in double existence depends on the vulgar belief, it is subject to any epistemic degradation resulting from the propensities that cause the vulgar belief.

For Hume, man arrives at monotheistic belief only after polytheistic belief in invisible intelligent power has been transformed by the propensity toward a system and anthropomorphized into a more robust polytheism. Only then does the propensity to adulate result in monotheistic belief. Thus, monotheism is subject (at the very least) to whatever epistemic difficulties face polytheism, including the negative consequences derived from the propensity to anthropomorphize.<sup>20</sup>

#### *4e. Hume's Attitudes Regarding the Propensities that Cause Religious Belief*

There is additional evidence that Hume thinks the propensities responsible for religious belief, taken together, are unlike custom and the other warrant-producing belief-forming mechanisms. He claims that the mechanisms that produce religious belief are known to produce other false beliefs: "the same principles naturally deify mortals, superior in power and courage, or

---

20 One might argue here that monotheistic belief may be warranted in some other way, via an argument, testimony, empirical evidence, or some direct propensity to believe in a designer. I briefly review evidence in Section 2 that Hume does not think monotheistic belief is justified by any argument, and I argue in my "Cleanthes's Propensity and Intelligent Design" that Hume does not think belief is warranted on the basis of a propensity to believe in a designer; however, further discussion of this issue is beyond the scope of this paper.

understanding, and produce hero-worship, together with fabulous history and mythological tradition, in all its wild and unaccountable forms” (NHR 5.9; 152). Given that he argues that the same propensities cause other problematic beliefs, we have reason to reject religious belief, just as we have reason to reject other mythological traditions.

Moreover, in the Introduction to the *Natural History*, Hume claims:

This preconception [invisible, intelligent power] springs not from an original instinct or primary impression of nature...first religious principles must be secondary; such as may easily be perverted by various accidents and causes, and whose operation too, in some cases, may, by an extraordinary concurrence of circumstances, be altogether prevented. (NHR Introduction.1; 134)

Thus Hume thinks that while the propensities that cause religious belief are common, they more strongly resemble those he opposes to custom—those that are rejected by philosophy: “neither unavoidable to mankind, nor necessary, or so much as useful in the conduct of life... [and] may easily be subverted by a due contrast and opposition” (T 1.3.4.1; 225). Since religious beliefs are unlike custom in this way, they ought to be rejected.

At this point, one might object that the causal explanation of the religious belief (or any belief) does not lead to rejection of religious belief (or any other belief), as this would be an instance of the genetic fallacy. Instead, they might argue that this generates only skepticism about the belief. For example, if I have a belief that it is 9.15am now because I made a guess, the fact that I formed that belief by guessing does not provide a ground for me to reject the belief that it is 9.15am, though I am skeptical about whether it actually is 9.15am now. Thus, the objector may claim that after learning the pernicious causes of religious belief, we have reason to temporarily suspend our belief; and then we must determine whether belief is justified in some other way--we must examine the evidence and arguments both in favor of and against religious belief.

However, in this particular case, the situation is different. In Hume’s view, religious belief is not justified on the basis of any cogent philosophical arguments (see Section 2, above). Moreover, the fact that Hume provides a causal explanation of religious belief is reason to think that it cannot be justified in some other way. If a belief is justified by observation, by philosophical argument, or by testimony, then the questions that motivate Hume to provide a genetic explanation do not arise. In general, while there will still be a natural cause for a belief, that cause will not normally affect the warrant of a justified belief. Only if the belief is not warranted on the basis of evidence or philosophical argument does the naturalistic explanation for belief play a role in its epistemic status. Hume rarely if ever examines the causal explanation of a belief that is justified; it is only *after* he has discovered that the belief is not reasonable (false, meaningless, or unjustified) that an



examination of its origin becomes especially pressing. Given that Hume provides a naturalistic explanation for religious belief, we have good reason to believe that Hume doesn't think it warranted by some other means.

Thus, for the case of religious belief, Hume's argument comes in two parts. First, in Parts IX and X of the first *Enquiry* and in the *Dialogues*, Hume contends that there is no cogent *argument* to justify religious belief. This parallels his claim in *Treatise* 1.3.6 that no cogent arguments support belief in the uniformity of nature. Second, in the *Natural History*, the propensities that cause religious belief confer negative epistemic status on religious belief. In this regard, the causes of religious belief are similar to the causes of the beliefs of the ancient philosophers—in substrata, sympathies, antipathies, and so forth. With the second part of Hume's argument, the parallelism with inductive beliefs—those caused by the “universal, permanent, and irresistible” principles of human nature—comes to an end. When we learn that the beliefs of the ancient philosophers, for example, are caused by a propensity Hume explicitly contrasts with custom, we ought to avoid or reject those beliefs. There is no need to reconsider the arguments in favor of those beliefs. Thus, Hume's causal explanation of religious belief provides grounds to *reject* religious belief.<sup>21</sup>

## 5. Conclusion

Hume's explanation of religious belief in the *Natural History* does not support the claim that religious belief is among the Humean natural beliefs.<sup>22</sup> In examining the details of the propensities responsible for religious belief, it is clear that religious beliefs (both polytheism and monotheism) have negative epistemic status. The propensities that cause religious belief are contrasted explicitly with custom and the other “permanent, irresistible, and universal” propensities of the imagination. Religious belief is produced in circumstances of fear and ignorance by belief-forming mechanisms not aimed at truth. Hume's description of the beliefs themselves, containing terms like “exaggerated, superstitious, absurd, and the like” provide further evidence that religious belief as described in the *Natural History* is unwarranted.

---

21 Rejecting religious belief does not mean that one may endorse the opposite. While Hume's attack on the arguments in favor of religious belief and his discussion about the causes of religious belief provide reason to reject religious belief, this does not give license to positively endorse claims such as “God does not exist.” In order to do that, one would have to have arguments, evidence, or reliable testimony for that claim (or it would have to be a natural belief based on a warrant-conducting propensity). Hume, however, provides us with no arguments or evidence or genetic description of atheism.

22 At this point, one might object that while the propensities that cause religious belief as described in the *Natural History* do not produce natural beliefs, there's another source: The *Dialogues*. Perhaps there's another form of religious belief, akin to “genuine theism” or “attenuated deism,” not justified on the basis of the argument from design, but instead caused by a universal propensity to believe in a designer triggered by the observation of apparent *telos* in nature. It's possible, then that *this* belief, with this cause, is a natural belief. I argue against this objection in my “Cleanthes's Propensity and Intelligent Design” (2011).

In the *Dialogues* and other texts, Hume argues against a foundation in reason for religious belief and concludes that religious belief is not warranted by testimony or cogent philosophical argument. In the *Natural History*, he provides what he thinks is the only other viable explanation for religious belief and shows that religious belief is not caused by warrant-generating psychological propensities. Rather, he claims that religious belief is based on psychological propensities (toward a system, to anthropomorphize, and to adulate) that generate unwarranted beliefs. The *Natural History* provides a crucial part of Hume's attack on religious belief, which, taken as a whole, suggests that we ought to reject religious belief in all of its forms.

## Bibliography

- Andre, Shane (1993) "Was Hume and Atheist?" *Hume Studies* 19, no. 1: 141–166.
- Arnold, N. Scott (1983) "Hume's Skepticism about Inductive Inference," *Journal of the History of Philosophy* 21, no. 1: 31–55.
- Badia Cabrera, M. A. (1995) "Hume's 'Natural History of Religion': Positive Science or Metaphysical Vision of Religion," in Tweyman (1995).
- Beauchamp, Tom L. and Mappes, Thomas A. (1975) "Is Hume Really a Sceptic about Induction?" *American Philosophical Quarterly* 12, no. 2: 119–29.
- Beauchamp, Tom L. and Rosenberg, Alexander (1981) *Hume and the Problem of Causation*, Oxford University Press: New York.
- Black, Tim and Gressis, Robert (2017) "True Religion in Hume's *Dialogues Concerning Natural Religion*," *British Journal for the History of Philosophy* 25, no. 2: 244–264.
- Broughton, Janet (1983) "Hume's Skepticism about Causal Inferences," *Pacific Philosophical Quarterly* 64, no. 1: 3–18.
- Butler, R. J. 1960. "Natural Belief and the Enigma of Hume," *Archiv fur Geschichte der Philosophie* 42 (no. 1): 73–100.
- Craig, Edward (1987) *The Mind of God and the Works of Man*, Clarendon: Oxford.
- Clifford, W. K. (1877) *Lectures and Essays*, Vol. 11, Macmillan: London.
- Connon, R.W. (1976) "The Naturalism of Hume Revisited," in Norton, Capaldi, and Robison.
- Costelloe, Timothy M. (2004) "'In every civilized community': Hume on Belief and the Demise of Religion," *International Journal for Philosophy or Religion* 55: 171–185.
- Falkenstein, Lorne (2003) "Hume's Project in 'The Natural History of Religion,'" *Religious Studies* 39: 1–21.

- Ferreira, Jamie M. (1995) "Hume's *Natural History*: Religion and 'Explanation,'" *Journal of the History of Philosophy* 33, no. 4: 593–612"
- Fogelin, Robert J. (1985) *Hume's Skepticism in the Treatise of Human Nature*, Routledge & Kegan Paul: London.
- Gaskin, J. C. A. (1974) "God Hume and Natural Belief," *Philosophy* 49, no. 189: 281–294.
- Gaskin, J. C. A. (1988) *Hume's Philosophy of Religion*, 2nd Edition, Macmillan: London.
- Harris, H.S. (1987) "The 'Naturalness' of Natural Religion," *Hume Studies* 13, no. 1: 1–29.
- Holden, Thomas (2010) *Specters of False Divinity: Hume's Moral Atheism*, Oxford University Press: Oxford.
- Holley, David M. (2002) "The role of Anthropomorphism Hume's critique of theism," *International Journal for Philosophy of Religion* 51: 83–99.
- Hume, David (1993) *Dialogues Concerning Natural Religion and Natural History of Religion*, edited by J. C. A. Gaskin, Oxford: New York.
- Hume, David (1947) *Dialogues Concerning Natural Religion*, 2nd edition, edited by Norman Kemp Smith, Thomas Nelson and Sons: Edinburgh.
- Hume, David (2000a) *An Enquiry Concerning Human Understanding*, edited by Tom L. Beauchamp, Clarendon: Oxford. 31
- Hume, David (2000b) *A Treatise of Human Nature*, edited by David Fate Norton and Mary J. Norton, Clarendon: Oxford.
- Hume, David (1978) *A Treatise of Human Nature*, edited by L. A. Selby-Bigge, revised by P. H. Nidditch, 2nd Edition, Clarendon: Oxford.
- Hume, David (2007) *The Natural History of Religion*, edited by Tom L. Beauchamp, Clarendon: Oxford.
- Hurlbutt, Robert (1985) *Hume, Newton and the Design Argument*, Lincoln: University of Nebraska Press.
- Garrett, Don (2012) "What's True about Hume's 'True Religion?'" *The Journal Of Scottish Philosophy* 10, no 2: 199–220.
- Goodnick, Liz (2011) "Cleanthes's Propensity and Intelligent Design," *The Modern Schoolman* 88, no. 3/4: 299–316.
- Jones, Peter (1972) "Hume's Two Concepts of God," *Philosophy* 47, no. 182: 322–333.
- Immerwahr, John (1984) "David Hume on Incompatible Religious Beliefs," *International Studies in Philosophy* 16: 25–34.
- Immerwahr, John (1996) "Hume's Aesthetic Theism," *Hume Studies* 22, no. 2: 325–337.

- Kail, P. J. E. (2007a) *Projection and Realism in Hume's Philosophy*, Oxford: New York.
- Kail, P. J. E. (2007b) "Understanding Hume's 'Natural History of Religion,'" *The Philosophical Quarterly* 57, no. 227: 190–211.
- Kemp Smith, ed. (1947) *Dialogues Concerning Natural Religion*, Thomas Nelson and Sons: Edinburgh.
- Kraal, Anders (2013) "Philo's Argument from Evil in Hume's *Dialogues X*: A Semantic Interpretation," *Sophia* 52: 573–592.
- Lemmens, Willem (2012) "Hume's Atheistic Agenda," *Bijdragen* 73, no. 3: 281–303.
- Lenz, John W. (1958) "Hume's Defense of Causal Inference," *Journal of the History of Ideas* 19, no. 4: 559–67.
- Loeb, Louis (2002) *Stability and Justification in Hume's 'Treatise'*, Oxford: New York.
- Logan, Beryl (1992) "The Irregular Argument in Hume's *Dialogues*," *Hume Studies* 18, no. 2: 483–500.
- Logan, Beryl (1993) *A Religion Without Talking: Religious Belief and Natural Belief in Hume's Philosophy of Religion* (Studies in European Thought Book 7), Peter Lang International Academic Publishers: New York
- Logan, Beryl (1996) "Why Hume Wasn't an Atheist: A Reply to Andre," *Hume Studies* 22, no. 1: 193–202.
- Fate Norton, David, Capaldi, Nicholas, Robison, Wade L., eds. (1979) *McGill Hume Studies*, Austin Hill Press: San Diego.
- Noxon, J. (1995) "In Defence of 'Hume's Agnosticism'" in Tweyman (1995).
- Penelhum, Terence (1979) "Hume's Skepticism and the Dialogues," in Norton, Capaldi, and Robison.
- Penelhum, Terence (1983) "Natural Belief and Religious Belief in Hume's Philosophy," *The Philosophical Quarterly* 33, no. 131: 166–181.
- Pike, Nelson, ed. (1970) *Dialogues Concerning Natural Religion*, Bobbs-Merrill Co.: Indianapolis.
- Plantinga, Alvin (2000) *Warranted Christian Belief*, Oxford: New York.
- Prado, C.G. (1981) "Hume and the God-Hypothesis," *Hume Studies* 7, no. 2: 154–163.
- Reich, Lou (1998) *Hume's Religious Naturalism*, University Press of America: Lanham, New York, Oxford.
- Russell, Paul (2008) *The Riddle of Hume's 'Treatise': Skepticism, Naturalism, and Irreligion*, Oxford: New York.

- Russell, Paul and Kraal, Anders (2017) "Hume on Religion" *Stanford Encyclopedia of Philosophy*; <https://plato.stanford.edu/entries/hume-religion/>
- Schmitt, Frederick F. (1992) *Knowledge and Belief*, Routledge: London.
- Spector, Jessica (2003) "Value in Fact: Naturalism and Normativity in Hume's Moral Psychology," *Journal of the History of Philosophy* 41, no. 2: 145–63.
- Tweyman, Stanley, ed. (1995) *David Hume: Critical Assessments*, vol. 5, Routledge: New York.
- Webb, Mark (1991) "The Argument of the Natural History," *Hume Studies* 18, no. 2: 141–159. Winters, Barbara (1979) "Hume on Reason," *Hume Studies* 5, no. 1: 20–35.
- Wolterstorff, Nicholas (1996) *John Locke and the Ethics of Belief*, Cambridge University Press: New York.



## AS “MEN OF SENSE”. GODWIN, BAROJA, BATESON AND HUME’S “OF NATIONAL CHARACTERS”

**Abstract:** *Men of sense, Hume says, condemn the extreme undistinguishing judgments concerning national characters; yet, he adds, they also allow that each nation has a national character or a peculiar set of resembling manners. Hume’s “Of national characters” was published at the end of 1748 in unclear circumstances, but it is still the object of several discussions for different reasons. William Godwin, Julio Caro Baroja and Gregory Bateson seem to refer to it, even though only the first two acknowledge it. Godwin uses it as a weapon to attack the climatic theory in the service of tyranny; Baroja as a sceptical solvent to destroy all mythical national character and real national prejudice; Bateson as a model to delineate an abstract frame for the research on national differences. Since, as Hume warns us, we run with avidity to give our evidence to what flatters our national prejudices and, as Mary Wollstonecraft denounces, we are eager to give a national character to every people, “Of National Characters” still provides us with acute and instructive remarks: to speak of national characters does not necessarily mean that we are speaking in favour of nationalism and against the individuals.*

**Keywords:** *David Hume, National characters, William Godwin, Julio Caro Baroja, Gregory Bateson*

“Hume is indeed full of acute remarks, or he would not be Hume. [...] it would be well if popular writers of the present day had emancipated themselves from the delusions which perplexed his unsurpassed keenness of vision. Perhaps the most instructive example of his method is the interesting Essay on ‘National Characters.’” (Leslie Stephen, *History of English Thought in the eighteenth century*, 1876)

“You made a remark about ‘national character’ that shocked me by its primitiveness [...] what is the use of studying philosophy [...] if it does not make you more conscientious than any... journalist in the use of the DANGEROUS phrases such people use for their own ends” (Ludwig Wittgenstein, 16 November 1944)

### The character of having no character<sup>1</sup>

“When you come back to England from any foreign country, you have immediately the sensation of breathing a different air”, George Orwell observes in 1941; but then “the vastness of England swallows you up,

1 I’m very grateful to María Blanco Gonzáles, Laura Nicolí, Alberto Mingardi, Gianluca Mori and Emanuele Ronchetti for their suggestions.

and you lose for a while your feeling that the whole nation has a single identifiable character”.<sup>2</sup> And Orwell asks: “Are we not 46 million individuals, all different?”.<sup>3</sup> National characteristics are “not easy to pin down” and often turn out to be “trivialities”;<sup>4</sup> yet England will still be England, “having the power to change out of recognition and yet remain the same”.<sup>5</sup>

Like Daniel Defoe (“Thus from a Mixture of all Kinds began, / That Het’rogeous Thing, *An Englishman*”),<sup>6</sup> in 1748 David Hume claimed that the English are remarkable for their “wonderful mixture of manners and characters”: the great liberty of every man “allows him to display the manners peculiar to him”, hence the English “have the least of a national character; unless this very singularity may pass for such” (E-NC, 207; cf. HOE V, 132).<sup>7</sup> Joseph Priestley recalls Hume’s judgment,<sup>8</sup> and Immanuel Kant advances his critical opinion, which is close to that of the criticized author. Hume is “mistaken”, because “affectation of a character is precisely the general character of the people”, which Kant depicts in his own way: “*arrogant rudeness*”.<sup>9</sup>

The character of having no character: While Gotthold Ephraim Lessing applied a similar formula to the Germans,<sup>10</sup> Georg Christoph Lichtenberg, who wanted to have “Hume for a coffee”, put it among his aphorisms and agreed with it.<sup>11</sup> What about Hume’s legacy and national characters?

### *The essay and its readers*

“I would rather tell you of a fine dissertation, where you give a much greater influence to moral than to physical causes”, Montesquieu wrote to Hume:

as far as I am able of judging, this subject is deeply treated, how difficult it is to be treated, and written with a master’s hand and filled up with very new ideas and reflections.<sup>12</sup>

2 G. Orwell, *Inside the Whale and Other Essays*, Penguin: London 1962, 64 (cf. *ivi*, 72).

3 *Ivi*, 64.

4 *Ivi*, 75.

5 *Ivi*, 90 (cf. *ivi*, 64).

6 D. Defoe, *The True-Born Englishman*, London: J. Roberts, 1716, 11.

7 “Men have no certain Characters; or if they have, ’tis that of having no constant, unalterable, distinguishing one” (*The Works Of Monsieur De La Bruyere*, 2 vols., London: E. Curll *et al.*, 1713, II, 252; cf. *ivi*, 83).

8 J. Priestley, *Lectures on History and General Policy*, Dublin: P. Byrne, 1788, 446–447.

9 I. Kant, *Anthropology, History, and Education*, edited by G. Zöllner and R.B. Loudon, Cambridge: Cambridge University Press, 2007, 407.

10 G.E. Lessing, *Selected Prose Works*, edited by E. Bell, London: G. Bell and Sons, 1890, 490.

11 G.C. Lichtenberg, *Aphorismen. Zweites Heft: 1772–1778*, edited by A. Leitzmann, Berlin: B. Behr, 1904, 135; on Hume and the English, cf. *ivi*, 133, 289.

12 *Oeuvres complètes de Montesquieu*, 3 vols., edited by M.A. Masson, Paris: Les éditions Nagel, 1950–1955, III, 1230 (transl. mine).



Montesquieu was slightly ironic: he knew that Hume was acquainted with the *Spirit of the Laws*, published one month before Hume’s essay; but these are but conjectures without evidence. Montesquieu’s is probably the first private review of Hume’s essay. Rejecting “altogether” the physical causes, Nicolas-Claude Thieriot remarks, Hume “goes beyond what he ought and even beyond his own opinion”, while Montesquieu “conceded too much to climate”;<sup>13</sup> his rejection, Jean-Bernard Mérian echoes, is “directly contrary” to Montesquieu.<sup>14</sup>

The topic of national characters was often debated in eighteenth-century Europe, especially in Scotland: Adam Ferguson’s *Essay* (1767), John Millar’s *Observations* (1771) and *Origin* (1779), Henry Home’s *Sketches* (1774), William Robertson’s *History* (1777), and James Dunbar’s *Essays* (1780). Here I will consider three exemplary uses of Hume’s essay: the first use is avowed and common in the eighteenth and nineteenth centuries; the second is equally avowed and common especially in the twentieth century; the third is merely conjectural and certainly uncommon in the twentieth century. In 1793 William Godwin followed Hume and attacked the explanation by physical causes, in 1970 Julio Caro Baroja appealed to him to destroy the myth of national character, and in 1942 Gregory Bateson could have been inspired by Hume in his positive treatment. Before analysing these uses, I will go back to the essay.

## Characters and prejudices: Hume’s double view

In 1734 Hume makes his first remark on national characters: “’tis with Nations as with particular Man, where one Trifle frequently serves more to discover the Character” (*Letters* 1: 21) – Lawrence Sterne will see the marks of national characters in some “nonsensical *minutiae*”.<sup>15</sup> In 1748, with Richard Steele, Hume observes that “nothing serves more [than travelling] to remove [national] Prejudices” (*Letters* 1: 126).<sup>16</sup> These prejudices, he maintained in 1739, are the effect of rash general rules: we establish it as a principle that the Irishmen cannot have wit, and we entertain such a prejudice against them “in spite of sense and reason” (T 1.3.13.7; SBN 146–147). As part of the “vulgar” we are “commonly guided” by that influence of general rules which is “destructive” of reasoning and rejected by “wise men” (T 1.3.13.12; SBN 150). Hume accounts for national characters from sympathy rather than climate:

13 [N.-C. Thieriot] “Sur le Caractère des Nations”, *Mercure de France*, Vol. 2 (1756): 21–49: 24 n. 2.

14 [J.-B. Mérian] “Sur le Caractère des Nations”, *Journal Encyclopédique*, Vol. 3 (1757): 26–51: 30.

15 [L. Sterne] *A Sentimental Journey Through France and Italy*, 2 vols., London: T. Beckett and P.A. De Hondt, 1768, I, 159.

16 Cf. [R. Steele] *The Guardian*, 2 vols., London: J. Tonson, 1714, Vol. 1, No. 34 (30 April 1713): 135–138: 134–135.

the “great uniformity [...] in the humours and turn of thinking of those of the same nation” must be ascribed to the “propensity we have to sympathize with others, and to receive by communication their inclinations and sentiments” (T 2.1.11.2; SBN 316–317).

Hume’s reflections in the *Treatise of Human Nature*, and his distinction between the “vulgar” and “men of sense” at the beginning of “Of National Characters”, remember us that there is no acceptance for absolute judgments. Men of sense reject all the vulgar undistinguishing judgments that “comprehend every individual under the same censure”; yet they allow that a certain nation has a certain character and that some qualities are more “probably” and “frequently” to be met with among the members of a nation than its neighbours, “though” the same men of sense will not forget the exceptions (E-NC, 197–198).<sup>17</sup>

Even Hume shows a prudent good sense. His language is full of “perhaps” and “may be”. He affects caution with some ambiguities (he is “inclined” to doubt, though “altogether”; he allows that the contrary opinion “may justly [...] seem probable”, though “at first sight”, E-NC, 200, 201), especially when he advances some controversial or embarrassing hypothesis as a suspicion (E-NC, 202 n.4, 208 n.10). With regard to the “promiscuous” characters of nations in temperate climates, he even uses a sceptical formula: “this however is certain, that [...] almost all the general observations [...] are found to be uncertain” (E-NC, 208). Against these observations, which appeal to physical causes and acknowledged authorities, Hume cautiously advances his own exceptions and moral causes (E-NC, 207; 209–210, 211, 213–215).

“Of national characters” was published at the end of November 1748, after six months abroad to the courts of Vienna and Turin and eight years of war, and achieved its final structure in 1753.<sup>18</sup> Hume believes in the existence of “national character”: it is a “uniform”, “common” and “peculiar” character consisting in a “peculiar set of manners”, in a “resemblance” or “similarity” of manners that are possessed by a people or nation and make it “distinguishable”. Yet, not all the qualities are always habitual and extended to the whole nation (NC, 197–99, 202–207).

Hume’s starting point is a question: if physical causes have a mighty influence on animals and plants, if animal characters are “derived” from their

17 Cf. A. Ferguson, *An Essay on the History of Civil Society*, edited by F. Oz-Salzberger, Cambridge: Cambridge University Press, 2001, 183).

18 In the 1753 edition Hume divides the first paragraph in two, turns the discursive footnote on courage into a paragraph, and makes several additions: two discursive footnotes on the character of animals and “negroes”; another discursive footnote on the small sects; a paragraph on revenge and the final sentence on the sacerdotal office in the footnote on priests; a sentence on the taciturnity of the Gauls according to Diodorus Siculus in the last footnote; finally, the references to Cicero and Quintilian in the text, Titus Livius, Quintus Curtius and Plutarch in the footnotes. We may conjecture that the 1748 version was written in a hurry (even though not necessarily in Turin after the reading of Montesquieu).

native climate, “why not the same with men?” (E-NC, 201–202; cf. E-NC, 207). By distinguishing national characters from prejudices, Hume establishes the legitimate use of the notion (E-NC, 197–198). With regard to their causes, he sides with those who explain national characters by moral rather than physical causes, and sets sympathy against climate (E-NC, 198–204). Then he considers the first objection from animals (in 1753 he expresses the “little suspicion” that animal characters do not depend on climate but on the different breeds and their rearing, E-NC, 202 n.4), and replies by appealing to sympathy, fixed moral causes and accidents (the prevailing disposition in the infancy of society, and the persons in authority at the first establishment of a government).<sup>19</sup> Seeking for proofs in history and geography he advances nine observations in favour of moral causes, especially government (E-NC, 204–207). Then he considers the second objection from animals and plants, where the influence of the degrees of heat and cold seems to be proved by the inferiority of those who live beyond the polar circles or between the tropics (in 1753 he is “apt to suspect” that the inferior character of the “*Negroes*” does not depend on climate but on some natural original difference between the breeds of men, E-NC, 208 n.10),<sup>20</sup> and advances a probable explanation from moral causes or indirect physical causes (E-NC, 207).

Here begins the second part of the essay, where Hume shows the fallacy of the general observations on national characters in temperate climates. He accounts for eight observations in favour of climate supported by authorities (Bacon, Berkeley, Bentivoglio and Temple) and shows that they are not universally valid (E-NC, 208–212). With regard to the distribution of the passion for alcohol in the North and love in the South, by way of mere hypothesis, he first accepts the seemingly probable explanation from heat and cold; then he offers his possible explanation from moral causes or indirect physical causes; then he calls into doubt the distribution; finally, he acknowledges the physical explanation of the distribution, but limits its consequences: climate can only have an influence on gross bodily organs, not on those fine organs on which the operations of the mind depend (E-NC, 213–215).

The character of a nation consists in a peculiar set of resembling manners that distinguish a particular “rank” of people: the “generality” or the “greater number” (E-NC, 197–199). Hume had already explained why a nation is

19 According to *A Dialogue* (1751) “chance has a great influence on national manners” (D 50; SBN, 340).

20 Our knowledge of Hume’s “racism” (a term unknown in his time) is inferior to the number of words on it. See for example E. Mazza, *Cavalli bianchi e pappagalli neri*, “I castelli di yale”, Vol. I, No. 2 (2013): 187–220; “Black Parrots”, The Hume Society’s 41<sup>st</sup> Annual Conference, Portland, Oregon, 2014, conference handbook; A.C. Willis, “The Impact of David Hume’s Thoughts about Race for His Stance on Slavery and His Concept of Religion,” *Hume Studies*, 42, 1–2 (2016): 213–239; A. Garrett and S. Sebastiani, “David Hume on Race,” in *The Oxford Handbook of Philosophy and Race*, edited by N. Zack, Oxford: Oxford University Press, 2017, 31–43.

more learned than others by appealing to its government, and maintained that the “multitude” will “certainly” be seized by the common passion, even though “many individuals may escape the contagion, and be ruled by passions peculiar to themselves” (E-RP, 115).<sup>21</sup> Likewise, here he maintains that in England every individual can display “the manners peculiar to him” (E-NC, 207).

Even though national characters depend on moral causes and accidents, government and sympathy, Hume makes some concessions to indirect physical causes (E-NC, 207, 213): southern people have a greater inclination to love because they go half-naked (direct moral cause) because of the “heat” (E-NC, 213).<sup>22</sup> We can even see the relics of that physical explanation Hume has just destroyed: in temperate climates blood is not so inflamed to “render” men jealous, but warm enough to “make” them to set a due value on the charm of women (E-NC, 214–215). In general, Hume reminds the reader two things: as the common character is spread by “sympathy or contagion of manners”, the closer the communication is, the greater is the similarity of manners (E-NC, 202–204); the “manners of a people change very considerably from one age to another” (E-NC, 205) or, as he adds in 1753, “national characters may alter very much” (E-NC, 213 n.17).

Hume leaves a complex legacy. As he had previously maintained (cf. T 2.3.1.10; SBN 403), national characters do exist but should not be confounded with prejudices. They consist in that peculiar set of resembling manners that distinguish a nation. They admit of exceptions and change in the course of time. They spread themselves by sympathy and communication. Their causes are not physical, but moral, even though climate may have an indirect influence (if it had a direct influence, it will insensibly affect only our bodily organs). Where characters do not depend on fixed moral causes, they depend either on accidents or an original natural difference. Finally, we must treat the matter with a sceptical attitude: we must limit ourselves to tendencies and probabilities and be cautious in our observations on characters in temperate climates, the extension and constancy of certain qualities, the natural distribution of certain passions.

From time to time, readers will receive or reject some elements of this legacy. Two dark footnotes have been the main target: Alexander Gerard and others attacked Hume on “priests”,<sup>23</sup> James Beattie and others on “negroes”.<sup>24</sup> Gerard denounced Hume’s contradiction: he “justly blames the

21 Cf. B.-L. De Muralt, *Lettres sur les Anglois et les François et sur les Voyages*, [London?] s.e., 1725, 175–176.

22 Cf. R. Susato, *Hume’s Sceptical Enlightenment*, Edinburgh: Edinburgh University Press, 2015, 52.

23 *Early Responses to Hume’s Moral, Literary and Political Writings*, edited by J. Fieser, 2 vols., Bristol: Thoemmes Continuum, 2005, II, 193–195; for Wallace’s, Shaw’s, and Ogilvie’s reaction, cf. *ivi*, 72, 351–352, 375–382.

24 *Ivi*, 276–277; for Swediaur’s and Kames’ reaction cf. *ivi*, 383–386; H. Home, *Lord Kames, Sketches of the History of Man*, 2 vols., London: W. Strahan and T. Cadell, 1774, I, 32–33.

undistinguished judgments of the vulgar”, he “justly observes, that all that can be asserted with truth is, that some particular qualities will be more frequently met with among some classes of people than others”; but he did not preserve “this necessary caution and delicacy in determining the character of the clergy”.<sup>25</sup>

“Philosophy – Hume declares – would render us entirely *Pyrrhonian*, were not nature too strong for it” (T.Abs.27; SBN 657). The phrase depicts his “very sceptical” philosophy as a compounded process that should not be reduced to one single element. The first paragraph of “Of National Characters” can be compared to this phrase: our nature would lead us into national prejudices, if it was not corrected by reflection; but the rejection of national prejudices does not entail that of national characters. The beginning of the essay is “a two-handed pot”:<sup>26</sup> “Men of sense condemn the undistinguishing judgments: Though at the same time, they allow, that each nation has a peculiar set of manners” (E-NC, 197). Readers commonly choose the denial of national prejudices. Some of them appeal to this denial with regard to the race question. In 1956 Geoffrey McKay Morant remarks that in ordinary conversation we speak “as if the collection were a single individual”:

the practice conveys the impression that the distinction between two groups compared is greater than it can be supposed to be if the case is examined in detail. This was Hume’s point when he remarked that the vulgar are apt to carry all national characters to extremes.<sup>27</sup>

In the 1974 edition of *Man’s most dangerous Myth* (1942), Ashley Montagu quotes Hume’s passage and concludes: “national and racial stereotypes mostly serve ill purposes”.<sup>28</sup>

Other readers appeal to the same passage with regard to the nationalist question. In 1955 Boyd Shafer quotes it and warns: a group cannot be described “as if it were a single man”: “the faulty reasoning is simple to demonstrate, though its effects are tragic”.<sup>29</sup> In 1972 he uses the passage to maintain that there is “no proof” that the citizens of one nation are more warlike than those of another: “individuals in each nation have been peaceful and warlike”.<sup>30</sup> According to the *Encyclopedia of Nationalism*, “as early as” the eighteenth century, Hume pointed out the road: “because of the many difficulties involved in studying the ramifications of stereotypes, some

25 *Early Responses*, II, 216.

26 Montaigne, *Les Essais*, 3 vols., edited by P. Villey, Paris: PUF, 1992, II, 581 (transl. mine).

27 G.M. Morant, *The Significance of Racial Differences*, in *The Race Question in Modern Science*, New York: UNESCO, 1956, 285–325: 288, 292–293.

28 A. Montagu, *Man’s Most Dangerous Myth: The Fallacy of Race*, 5<sup>th</sup> edn., Oxford: Oxford University Press, 1975, 284.

29 B.C. Shafer, *Nationalism. Myth and Reality*, New York: Harcourt, Brace & World, 1955, 332.

30 B.C. Shafer, *Faces of Nationalism. New Realities and Old Myths*, New York: Harcourt Brace Jovanovich, 1972, 339.

scholars believe that it is advisable to maintain them in myth”.<sup>31</sup> Let us now consider the three exemplary uses of Hume’s essay.

### “The judicious collections of Hume”: Godwin and the moral causes of freedom

The characters of passions and minds are different in different climates, and laws must have references to them: this doctrine, Reverend David Williams says in 1789, was the “principal distinction” of Montesquieu’s fame: the inferences from “physical effects” to political institutions are “so plausible” that they have been “generally adopted” and to dispute them is “a species of heresy”.<sup>32</sup> Hume “added to his numerous offences against popular opinion a general denial of the doctrine”,<sup>33</sup> and Williams advises his students: those who “mean to oppose the sentiments of Montesquieu, should not overlook those of Mr. Hume”.<sup>34</sup> Montesquieu maintained that “the cowardice of the peoples of hot climates has almost always made them slaves”;<sup>35</sup> in the *Enquiry Concerning Political Justice* (1793) Godwin replies to those who have “affirmed to be ‘impossible to establish a free government in certain warm and effeminate climates’” (PJ, 43):<sup>36</sup> “my original conception proceeded on a feeling of the imperfections and errors of Montesquieu”.<sup>37</sup>

Since the influence of climate could be an objection to his principles, Godwin calls it into question (PJ, 43). Chapter VII is divided in two parts: “Of moral and physical causes” and “Of national characters”. Following Williams’ advice, Godwin contrasts Montesquieu’s alleged facts with Hume’s exceptions. He begins the second part with a traditional comparison: “as is the character of the individual, so may we expect to find it with nations” (PJ, 42). According to Hume “a nation is nothing but a collection of individuals” (E-NC, 198), and Godwin echoes: “society is nothing more than an aggregation of individuals (PJ, 57, 71)”. Later he will both extend and restrained his view: “the universe is no more than a collection of individuals”, but “no man stands alone”.<sup>38</sup>

31 L.L. Snyder, *Encyclopedia of Nationalism*, New York: Paragon House, 1990, 376.

32 D. Williams, *Lectures on Political Principles*, London: J. Bell, 1789, 263–264.

33 Ivi, 264.

34 Ivi, 264.

35 Montesquieu, *The Spirit of the Laws*, edited by A.M. Cohler *et al.*, Cambridge: Cambridge University Press, 1989, 278. The assertion is attacked by Kames (*Sketches*, I, 31).

36 W. Godwin, *An Enquiry concerning Political Justice*, edited by M. Philp, Oxford: Oxford University Press, 2013 (hereafter PJ). On Hume and Godwin, see Susato, *Hume’s Sceptical Enlightenment*, 52–53.

37 C.K. Paul, *William Godwin: his Friends and Contemporaries*, 2 vols., London: H.S. King, 1876, I, 67, cf. *ivi*, 71.

38 W. Godwin, *An Enquiry concerning Political Justice*, 2 vols., London: G.G.J. and J. Robinson, 1796<sup>2</sup>, I, 203, 442.

Following Hume, Godwin maintains that the operations of law and political institution are “important”, those of climate are “trifling” (PJ, 42; cf. E-NC, 198, 200, 204). Like Hume and Adam Smith, he also claims that “particular professions [...] must always operate to the production of a particular character” (PJ, 42; cf. E-NC, 198–201),<sup>39</sup> and that, “in some instances”, some particular moral causes may render some men “superior” to their professional character (PJ, 42). Yet, he concludes with Hume, “exclusively of such exceptions, priests of all religions [...] have a striking similarity of manners” (PJ, 42; cf. E-NC, 198–199).<sup>40</sup> What is relevant for Godwin is that “free men in whatever country will be firm, vigorous and spirited in proportion to their freedom” (PJ, 43). To show how unreasonable is the assertion that we can’t establish a free government in a warm country, Godwin accounts for the process that establishes it in any country: to recommend the advantages of freedom to people, we must “inform” their understanding, rather than inquiring whether they are native of a “favourable” climate (PJ, 43). It is a question of truth, and the causes that suspend its progress arise “not from climate, but from [...] despotic sovereigns”: climate will not prevent the majority from embracing “the obvious means of their happiness” (PJ, 44).

To be acknowledged as true, Godwin’s theory needs to be confirmed by the history of mankind, where “the inhabitants of neighbouring provinces in different states, [are] widely discriminated by the influence of government, and little assimilated by the resemblance of climate” (PJ, 44; E-NC, 204). The majority of Godwin’s instances are “taken from Hume’s *Essay on National Characters*, where this subject is treated with much ability” (PJ, 46 n.1); they are “abridged from the judicious collections of Hume upon the subject”.<sup>41</sup> Godwin repeats almost verbatim (with some personal embellishments) the last edition of Hume’s essay<sup>42</sup>. Hume argues that “the Languedocians and Gascons are the gayest people in France; but whenever you pass the Pyrenees,

39 Cf. A. Smith, *The Theory of Moral Sentiments*, edited by D.D. Raphael and A.L. Macfie, Indianapolis: Liberty Fund, 1984, 202–203.

40 Godwin’s remarks on priests are drawn from Hume (E-NC, 199–201 n.3) and, like Hume, he admits of exceptions (E-NC, 200–201 n.3; PJ, 42).

41 Godwin, *Political Justice*, 1796, I, 100. In 1796 chapter VII (“Of the Objection to these Principles from the Influence of Climate”) is turned into chapter VI “Of the Influence of Climate” (ivi, I, 90): Godwin maintains part II (“Of national Characters”), omits part I (“Of moral and physical causes”), throws the rainy day and indigestion examples into another chapter (ivi, I, 79), adds some new passages (ivi, I, 96–97, 100–101) and cuts some old ones (ivi, I, 46–47), like that on priests (ivi, I, 42–43).

42 Godwin’s sale catalog lists Hume’s *Essays and Treatises on Several Subjects*, 2 vols., London: T. Cadell *et al.*, 1788; his diary has two entries “Hume on Nat. Characters”: 7 and 15 March 1792 (godwindiary.bodleian.ox.ac.uk). Godwin echoes Hume: “it is easy to obtain any thing of the negroes, even their wives and children” (PJ, 46). In 1772 Hume still claims that they sell “not only their parents, but their wives and mistresses” (D. Hume, *Essays and Treatises on several Subjects*, 2 vols., London, T. Cadell, 1772, I, 222); later he turns “parents” into “children” (E-NC, 214).

you are among Spaniards” (E-NC, 204), and Godwin, like Priestley,<sup>43</sup> echoes: “the Gascons are the gayest people in all France; but the moment we pass the Pyrenees, we find the serious and saturnine character of the Spaniard” (PJ, 44; cf. E-NC, 208).<sup>44</sup> Godwin also recalls Hume’s exceptions and remarks. Hume argues that “it would have been juster to have said, that most conquests are made by poverty and want upon plenty and riches” (E-NC, 201), and Godwin echoes: “it would have been truer to say that conquest is usually made by poverty upon plenty” (PJ, 45).<sup>45</sup> Sometimes Godwin makes a concession to Montesquieu (“to the extent that, in each nation, one of these causes acts more forcefully, the others yield to it”):<sup>46</sup> “physical causes have already appeared to be powerful, till moral ones can be brought into operation” (PJ, 46); but he remains convinced that the “result” of Hume’s reasonings is “of the utmost importance to him who speculates upon principles of government” (PJ, 46).

Godwin agrees with Hume that climate, if it operates, “may operate upon the grosser particles of our frame, not [...] those finer organs upon which the operations of intellect depend” (PJ, 46; cf. E-NC, 215). In the first part he distinguishes between those causes that operate directly on the mind as “subject of reasoning”, and “inducements to action”, and those that operate indirectly by rendering the animal frame gay or sluggish, and must be regarded as “inefficient” and “trivial” – the “meanest branch of human constitution” to which many thinkers ascribe “so much” (PJ, 37). Unlike bodily indisposition, which operates on the mind as a physical cause, corporal punishment operates as a moral cause: “it influences our conduct, only as it is reflected upon by the understanding, and converted into a motive of action” (PJ, 37–38).

To those who assert that a rainy day makes us cowards and indigestion unable to think (PJ, 38–39), Godwin replies: physical indisposition becomes “formidable” only if seconded by the consent of the mind; the communication with the material universe is “at mercy” of our choice; the inability of the understanding exists “only in the degree in which it is deliberately preferred” (PJ, 39). In the name of the association of ideas, Godwin makes his last attack:

If men were principally governed by external circumstances such as that of atmosphere, their characters and actions would be much alike. [...]. Every thing that [...] permanently distinguishes the character of

---

43 J. Priestley, *Lectures on History and General Policy*, Dublin: P. Byrne, 1788, 447; cf. ivi, 444–447.

44 Godwin recalls Hume’s observations against climate: on Athenians and Thebans (E-NC, 204; PJ, 44), Gascons and Spaniards (E-NC, 204; PJ, 44), Jews and Armenians (E-NC, 205 and n. 7; PJ, 45), Turks and Greeks (E-NC, 205; PJ, 45), Greeks, Italians and French (E-NC, 206; PJ, 45), and Gauls (E-NC, 213 n. 17; PJ, 45).

45 See PJ, 45–46; E-NC, 205, 207, 208, 213, 214, 215.

46 Montesquieu, *The Spirit of the Laws*, 310.



one man from that of another, is to be traced to the association of ideas. But association is of the nature of reasoning. [...]. It is thus that one man appears courageous and another cowardly [...] under the same or nearly the same external circumstances (PJ, 39).

Pain occurs “too seldom” to be mistaken for the efficient principle of character; the system which “determines” our proceedings springs from a different source and “returns” when the pain has subsided (PJ, 39–40).

Where our conduct is directed “merely” by external impressions, we resemble animals (PJ, 40). Like Hume, Godwin considers the mind in its “infancy”, where it is “nearly” the creature of contingencies, as the supporters of physical causes claim; yet, he adds, the farther it advances the more it individualizes and acquires personal habits. This progress shows the influence of reflection:

Physical causes, though of some consequence in the history of man, sink into nothing, when compared with the great and inexpressible operations of reflection. They are the prejudices we conceive or the judgments we form [...] that constitute the true basis of distinction between man and man. The difference between savage and savage indeed, in the first generation of the human species and in perfect solitude, can only be ascribed to the different impressions made upon their sense. But this difference would be almost imperceptible (PJ, 41).

In the early age men lived in solitude and physical causes had an influence. This is Godwin’s answer to the analogical argument from “the considerable effects that physical causes appear to produce upon brutes” (PJ, 41). Like Hume, he observes that “the races of animals perhaps never degenerate, if carefully cultivated”, but “we have no security against the wisest philosopher’s begetting a dunce” (PJ 41) (Hume says “a coxcomb may beget a philosopher”, NC-215). Godwin repeats Hume’s objection from animals:

If [...] climate were principally concerned in forming the characters of nations, we might expect to find heat and cold producing an extraordinary effect upon men, as they do upon plants and inferior animals (PJ, 45, cf. E-NC, 200–202, 207).

Godwin limits the consequences of this objection: the existence of physical causes “cannot be controverted” in animals, where they are “left almost alone”, but in men “their efficacy is swallowed up in the superior importance of reflection and science”: in a “race of negroes”, which lie in solitude from their infancy, the operation of breed “might perhaps” be rendered as “conspicuous” as in the different classes of horses; yet, in a situation of communication, “all parallel” would cease (EPJ, 41; cf. E-NC, 200–202, 208 n. 10).

The more the individual is civilized, the greater is the action of reflection (as Montesquieu put it, “nature and climate almost alone dominate savages”),<sup>47</sup> and the changes of character do not depend on climate:

He that would change the character of the individual, would miserably misapply his efforts, if he principally sought to effect this purpose by the operation of heat and cold, dryness and moisture upon the animal frame. The true instruments of moral influence, are desire and aversion, punishment and reward, the exhibition of general truth (PJ 41–42).

The influence of physical causes is a “blind and capricious” principle that taints all our conclusions on the “advantages of liberty”: some writers have concluded that the corruptions of despotism and the usurpations of aristocracy were “congenial to certain ages and divisions of the world, and under proper limitations entitled to our approbation” (PJ, 46). According to them climate opposes “a palpable barrier to the political improvement” and renders “impossible” the introduction of liberal principles; yet, “truth when properly displayed [is] omnipotent”,<sup>48</sup> and the “correspondence between national character and national government will be found [...] to arise out of the latter” (PJ, 104). No state or period of mankind can render men incapable of exercising their reason or make it necessary to hold them in a condition of pupillage: “the real enemies of liberty in any country are [...] those higher orders who profit by a contrary system” (PJ, 47).

In order to change, we must infuse the just views into a number of educated persons and give the people their guides. The error is not tolerating the worst forms of government for a time, but supposing the change impracticable and not pursuing it. Godwin’s target is climate in the service of tyranny; and Hume serves to dismantle the political use of the climatic hypothesis, which is found “incapable of holding out against a moment’s serious reflection” (PJ, 46). Hume taught him that his reasonings on liberty are confirmed by “general experience as to the comparative *inefficacy of climate*, and the superior influence of circumstances, political and social”.<sup>49</sup>

### “All are eager to give a national character”: a digression on Wollstonecraft’s caution

In 1788 Mary Wollstonecraft reviewed the *Essay on the Causes of the Variety* of the minister Samuel Stanhope Smith, who had advanced his *Strictures on Lord Kaims’s Discourse*.<sup>50</sup> Kames had directly attacked

47 Ivi, 310.

48 Godwin, *Political Justice*, 1796<sup>2</sup>, I, 96.

49 Ivi, 100.

50 M. Wollstonecraft, “Review of *An Essay on the Causes*”, *The Analytical Review*, Vol. 2, art. 7 (1788): 431–439; S.S. Smith, *Strictures on Lord Kaims’s Discourse on the «Original Diversity of Mankind»*, in *An Essay on the Causes of the Variety of Complexion and Figure*

Montesquieu: his “plausible” natural causes “are contradicted by stubborn facts”.<sup>51</sup> Neither tempers nor talents depend “much” on climate.<sup>52</sup> Smith is convinced that the doctrine of “one” race “illustrates the power of physical causes”.<sup>53</sup> Wollstonecraft shares this conviction: if Smith’s arguments are “conclusive”, then “vague conjectures”, which shake our confidence in the validity of the Mosaical account”, will melt before them.<sup>54</sup> Different external circumstances, such as “the situation of the country, forms of government, religious opinions, &c.”, Wollstonecraft observes recalling Hume (E-NC, 198, 207), have been “traced by the ablest politicians as the main causes of distinct national characters”.<sup>55</sup> To account for this variety Smith refers to “climate” and the “state of society”, and Wollstonecraft embraces this combination, but finds “more interesting” the appeal to the effects of society and education.<sup>56</sup>

1796 is the year of the second meeting between William and Mary, and of the publication of the second edition of Godwin’s *Political Justice* and Wollstonecraft’s *Letters*. Like Rousseau, she is convinced that in the capital “less of a national character is to be found than in the remote parts of the country” (L, 15),<sup>57</sup> where the separated habitations “allow the difference of climate to have its natural effect” (L, 25). However, she is the partisan of the indirect influence of climate (cf. L, 29–30, 22), when she does not directly appeal to moral causes: if the Swedish children do not have the graces of their age, this is “much more owing to the ignorance of mothers than to the rudeness of the climate” (L, 22; cf. 14, 23, 32, 41, 97). Sometimes even moral causes are not able to offer an exhaustive account (L, 23, 101). Yet, most travel writers remark only the differences “inspired by the climate” rather than the “numerous and unstable” varieties that the forms of government and religion “produce”; “all are eager to give a national character; which is rarely just, because they do not discriminate the natural from the acquired difference” (L, 33). Only the eighteenth-century “spirit of enquiry”, she argues, will “destroy the factitious national characters which have been supposed permanent, thought only rendered so by the permanency of ignorance” (L, 33).

---

*in the Human Species*, Philadelphia: R. Aitken, 1787, 1–31; Kames, *Sketches*, I, 1–43. Cf. S. Juengel, “Countenancing History: Mary Wollstonecraft, Samuel Stanhope Smith, and Enlightenment Racial Science”, *EKH*, Vol. 68, No. 4 (2001): 897–927.

51 Kames, *Sketches*, I, 31 (cf. E-NC, 204).

52 H. Home, Lord Kames, *Sketches of the History of Man*, 4 vols., 2<sup>nd</sup> edn., London: W. Strahan and T. Cadell, 1778, I, I, 62.

53 Wollstonecraft, “Review of *An Essay*”, 437; Smith, *Essay*, 165.

54 Wollstonecraft, “Review of *An Essay*”, 431–432, 439; Smith, *Essay*, 143, 201–203.

55 *Ivi*, 431.

56 *Ivi*, 433–434; cf. *ivi*, 436; Smith, *Essay*, 122.

57 M. Wollstonecraft, *Letters written during a short Residence in Sweden, Norway, and Denmark*, Oxford: Oxford University Press, 2009 (hereafter L); J.-J. Rousseau, *Emile or On Education*, edited by A. Bloom, New York: Basic Books, 1979, 468–469.

Unlike Hume and Godwin (E-NC, 207; PJ, 45), and Like Beattie,<sup>58</sup> Wollstonecraft denounces those who forget to consider that faculties remain “obtuse”, where not “sharpened” by self-interest and necessity, and passions are “weak”, where the necessities of life are “too hardly or too easily obtained” (L, 33):<sup>59</sup> “A people have been characterized as stupid by nature; what a paradox!”. Slaves have “no object to stimulate industry”, and some other peoples have “no aptitude” for arts and sciences because “the progress of improvement had not reached that stage which produces them” (L, 33).

Unlike Hume, she distinguishes between manners and national character (L, 107), and maintains that the former may shape the latter: the character of the French “is, perhaps, more formed” by their theatrical amusements than is generally imagined.<sup>60</sup> Since the energy of thinking proceeds “either from our education or manner of living”, the French frivolity “may be accounted for, without taking refuge in the old hiding place of ignorance: occult causes”.<sup>61</sup> Yet, like Hume, she maintains that “the character of every man is, in some degree, formed by his profession”,<sup>62</sup> and sometimes “the professional had indeed effaced the national character” (L, 107); that government exerts a strong influence, and public happiness “ought not to have been expected, before an alteration in the national character seconded the new system of government”.<sup>63</sup>

Did Mary Wollstonecraft – who in other matters was “rather inclined to coincide in opinion with Hume”<sup>64</sup> and to quote his *History and Essays* –<sup>65</sup> ever read Hume’s “Of National Characters”? We do not know it, but we know that she read Godwin’s *Enquiry* and that Godwin read Hume’s essay.

### “A threatening and dangerous myth”: Baroja and the vulgar

“All speaking of ‘national character’ is a mythical activity” (M, 34), declares Caro Baroja in *The Myth of National Character* (1970).<sup>66</sup> There are some precedents, besides Max Weber’s general sentence that “the appeal to

58 *Early Responses*, 277.

59 M. Wollstonecraft, *A Vindication of the Rights of Men. A Vindication of the Rights of Women. An Historical and Moral View of the French Revolution*, Oxford: Oxford University Press, 1999, 120.

60 Ivi, 298, 361.

61 Ivi, 365.

62 Wollstonecraft, *Vindication*, 82.

63 Wollstonecraft, *Vindication*, 343.

64 M. Wollstonecraft, “Review of *A sonnet*”, *The Analytical Review*, Vol. 7, art. 22 (1790): 304.

65 Wollstonecraft, *A Vindication*, 9, 124–125.

66 J.C. Baroja, *El mito del carácter nacional*, Madrid: Caro Raggio, 2004 (1970<sup>1</sup>) (hereafter M; all translations are mine); cf. E-S. Unamuno, *Weaving Identities in the Spanish Context*, in *Identity*, 2 vols. edited by M.M. Baptista, Coimbra: Grácio Editor, 2015, I, 31–49.

national character is generally a mere confession of ignorance”.<sup>67</sup> In 1944 Ludwig von Mises had attacked “the Fallacy of the Concept”: “there is no such thing as a stable national character”.<sup>68</sup> This arbitrary idea derives from a judgment which “omits all unpleasant facts contradicting the preconceived dogma”.<sup>69</sup> There is no “permanency”, no “stable character of which all its ideas and actions are the outcome”: the “generalization of features discovered in various individuals” is the result of an “ill-considered induction from an insufficient number of ill-assorted samples”.<sup>70</sup> To explain Nazism as the “outcome of the German national character”, and its “inherent tendency”, is a “vicious circle”.<sup>71</sup>

In 1940 Henry Hamilton Fyfe had published *The Illusion of National Character*, which was to be revised and shortened two years after Mises’ essay. When we have moved among many different nationalities, Fyfe argues, we come to know that “there are no ‘national character types’”, and we must remove “the most dangerous of the myths which in our time prevent the nations from settling down to make the best of life”.<sup>72</sup> This harmful illusion springs from the “ignorant belief in the differences between people inhabiting different territories”.<sup>73</sup> The target of Fyfe’s assault is the union of nation and honour:

If we admit national character, thus recognizing the existence of a certain mystical entity, the nation, we must also admit national honour [...]. But if we reject national character as an illusion, if we say that nations are masses of people who happen to have been born under the same Government, if we deny that any mystical bond unites them [...], then we must refuse to believe there can be any such thing as national honour.<sup>74</sup>

Since the interests of the individuals are not opposed to those of the nation, and mankind are beginning to realize their common interest, “the chief obstacle in the way of that realization is the illusion of national character”.<sup>75</sup>

In the name of historical method applied to anthropology, Baroja starts his campaign against national character and every abstract and synchronic

---

67 M. Weber, *The Protestant Ethic and the Spirit of Capitalism*, London: Routledge, 1992, 47.

68 L. von Mises, *Omnipotent Government*, New Haven: Yale University Press, 1945 (1944<sup>1</sup>), 234.

69 Ivi, 232.

70 Ivi, 231.

71 Ivi, 231–232, 234. We can correctly speak of the “a nation’s mentality” (“of the majority” and “subject to change”) “at a certain historical epoch” (ivi, 233).

72 H.H. Fyfe, *The Illusion of National Character*, London: Watts & Co., 1940, 21 (cf. ivi, 2). Fyfe Humeably underlines exceptions and changes (ivi, 34, 69, 103).

73 Fyfe, *The Illusion of National Character*, 1940, 2, 147; cf. ivi, 24, 52, 59, 79, 90, 130, 135.

74 Ivi, 125.

75 Ivi, 157.

anthropological study. His essay has an eloquent English epigraph, which is the first sentence of Hume's essay: "The vulgar are very apt to carry all national characters to extremes" (cf. E-NC, 197).<sup>76</sup> Holding one Humean handle Baroja attacks those north-American anthropologists who, in order to "nationalize" their science, produced a "considerable quantity" of works (M, 33), which are but failed attempts between science and "pleasant literature", being scarcely scientific, literary and pleasant (M, 33-34). They are the springs of "heavy sophisms, eternal common places... and painful constraints". National character is a false harmful myth:

He who speaks or chats adapts himself to a more or less elaborate tradition that is not based on scientifically observed or observable facts; and this tradition tends to give somewhat popular explanations and in fact changes itself much more than is believed or asserted. Myth is advantageous or disadvantageous according to those who elaborate or use it, and it may degenerate into real mania (M, 34).

One of Baroja's targets are the *Ideas of the Spaniards of the XVII Century* (1928) by Miguel Herrero García, which were enlarged four years before Baroja's essay and reflect the ideology of an erudite right-winger who thinks that the origin of what the one hundred percent Spanish (i.e., the Castilians) are and, "must think of themselves", lies in the middle of Sixteenth Century (M, 49-50). Baroja's other target is fascist literature. In the twentieth century the international conflict (Spanish against French) became also national (Catalans against Basques), and the "nationalistic thesis of nationalities" opposed itself to the "nationalistic thesis of State". The fascist literature, which appealed to the true "spanishness" or "italianness", was a reaction to the new national conflict. Between 1921 and 1945 the European fascist countries tried to fix the national characters and convert patriotism into monopoly and profit, and the American democrats rediscovered national characters:

What a sad discovery! What a poor discovery! What a miserable dialectical game! For we can frequently observe that indeed the authors of the most irritating judgments do not have great importance and, on the contrary, no value is given to the objective or positive judgments (M, 82).

The harmful myth must be unmasked, and Baroja goes back to the epigraph:

The myth of national character is a threatening and dangerous myth, like many myths of the pagan antiquity; yet, perhaps it has neither their majesty nor their depth. It makes councilmen speak much and badly, and Hume was right when he said that the "vulgar" carry its concepts

---

76 Cf. J.C. Baroja, *El Señor Inquisidor y otras vidas por oficio*, Madrid: Alianza Editorial, 1968, 74; J.C. Baroja, *De los Arquetipos y Leyendas*, Madrid: Ediciones ISTMO, 1991, 105, 130 n.217.

to extremes, and by “vulgar” today we mean many people who do not believe to be part of the vulgar (M, 82–83).<sup>77</sup>

Baroja identifies all his targets: the ascriptions of a “global character” to an entire people (“all the Spanish”), the “dogmatic” theories on their origin and way of being (M, 36, 72), and the tendency, often attending international conflicts, to characterize peoples in a “permanent way” (M 45, 47, 72, 75–77, 86 n.4, 87–88 n.6, 94 n.19). He attacks all those who, from a nationalistic or localistic, aggressive or defensive perspective, ascribe “permanent traits” to different peoples: the only permanent thing that can be found is the “will” to ascribe national characters as founded on supposed scientific facts (M, 79; cf. 54). Baroja unmasks all these clichés, contradictions, decompositions of national into regional characters, the stereotypical images that each nation produces of itself, where all defects are little and virtues are cardinal (M, 48), and the negative generalizations she produces of foreigners (M, 75).

In this “gallery of horror”, where Mead finds a place with her essay “National Character” (1953) and Bateson is never mentioned (M, 85 n.2), Hume is urged as the antidote. He is the sceptical myth destroyer, who presides over the beginning and the end of Baroja’s discussion. Hume’s “wonderful” essay, which Baroja dates 1742 (even a learned Spanish has his careless moments), appears to him “full of exceptions, reserves, fine observations which are contrary to any dogmatism in this matter”:

[Hume] points out the difference of general context between the eighteenth-century English society and that of the preceding century, discusses the influence of climate, insinuates that the *professional* character is perhaps more dominant than the *national* character, and finally delivers many other witty thoughts (M, 58).<sup>78</sup>

“Nevertheless”, Baroja acknowledges, Hume “ascribes to the Spanish, as a whole, a grave and sad character, and judge them decadent with respect to what made them famous in the preceding times, the soldier’s profession” (M, 58, 102 n. 44; cf. E-NC, 198, 204, 206, 208). Yet this ascription does not diminish Hume’s merit:

Everything the father of modern scepticism says is curious and useful today as well as in his time. His observations concerning the southern peoples compared to the northern, the little meaning of courage in the collective life, the constant variability of attitudes and knowledges are as many starting points for those who want to give the proper value

77 Baroja is evoked in the discussion of Kant’s answer (“a national character is not a mere chimera”) to the sceptical challenge (national character is “merely a myth”) (R.B. Louden, *Kant’s Human Being*, Oxford: Oxford University Press, 2011, 156, 199 n.16).

78 Baroja refers to D. Hume, *Essays and Treatises on Several Subjects* (2 vols., London: A. Millar, 1764) and following the frontispiece he maintains that it “appeared in 1742” (M, 101–102 n.44).

to myths without being drawn by the force of the above-mentioned common places that, as Hume points out in the sentence quoted at the beginning of these pages, make a strong impression on the vulgar: a vulgar among whom there are writers, ministers or generals... (M, 58).

Baroja attributes too much to his favourite Hume and shows some inaccuracies. Hume is more interested in the causes of characters than their content (E-NC, 198); he only reports the observations on the nature of the northern and southern peoples, and often disagrees with them (E-NC, 213–214); he does not insinuate that the professional character is more dominant than the national one (Wollstonecraft says that the professional may efface the national character), but maintains that moral causes fix the professional characters and alter the natural personal character, and the professional character does not prevail in every instance over the personal character but always predominates with the greater number (E-NC, 198–199);<sup>79</sup> he does not say that courage has little meaning, but that it is exerted only at intervals by a minority (E-NC, 212); finally, Hume does not merely discuss the influence of climate, but rejects it in favour of moral causes (E-NC, 198, 200, 203–204).

Baroja distinguishes Hume's view from Montesquieu's and Voltaire's "generalizing pretensions" (M, 59). Sometimes, he remarks, the impressions communicated by men of genius are the "most gloomy" (M, 102, n. 45), and their generalizations are devoid of any foundation:

the ignorance of every age, as well as its knowledge or more than it, helped to found its characterizations [...]. The possibility to err in general is parallel to that of deceive oneself, even frequently, in particular judgments: in the first case errors are considered as "scientific", in the other they are rejected as mere "judgments of value" (M, 59).

Baroja's radical reading of Hume's essay is founded on its beginning: "The vulgar are apt to carry all national characters to extremes [...] Men of sense condemn these undistinguishing judgments" (E-NC, 197). Yet, according to Hume, prudent men of sense also allow that each nation has her own character (with its exceptions). Baroja reduces national characters to national prejudices. Hume certainly insists on the "constant variability" of characters (they are variable as they depend on variable moral causes): he does not limit himself to saying that the "manners" of a people change very considerably, but adds that "national characters may alter very much" (E-NC, 205, 213 n.17). And he can say so because – as Ernest Barker acknowledged in 1948 –<sup>80</sup> he allows the existence of national characters. They can be dangerous, when carried to extremes, but are not mere harmful myths.

79 Cf. Gerard in *Early Responses*, II, 211.

80 Even Hume "assumes, as a matter of obvious common sense, that there *are* national characters" and his "only scepticism" is about the influence of climate upon them (E.



The “father of modern scepticism”, as Baroja calls him, does not treat national characters as other modern philosophical notions. He does not say that philosophy cannot justify their existence and, when consistently pursued, leads us into scepticism with regard to them; that we can only account for this belief upon the suppositions of the vulgar (all of us, at one time or another) and find its origin in the principles of our mind. The existence of national characters is something that we must take for granted as well as their causes and reductions to extremes.

### “A sort of resonance”: Bateson, Mead, and national morale

Thirty years before Baroja, Bateson grabbed the other Humean handle. We do not know whether Bateson ever read “Of national characters”. It is possible, not probable; but there are some analogies. The Select Society of Edinburgh (Hume was a founding member) was formed in 1754 to discover the most effectual methods of promoting the good of the country;<sup>81</sup> one of the much-debated question was “Whether the difference of national characters be chiefly owing to the nature of different climates, or to moral and political causes?”<sup>82</sup> In 1942, as a member of the Committee for National Morale, organized by Arthur Upham Pope in 1940, Bateson published “Morale and National Character”. Here he outlined a theoretical structure and certain “useful” formulas which are based on the assumption that “people will respond most energetically when the context is structured to appeal to their habitual patterns of reaction” (MNC, 89).<sup>83</sup>

At the end of 1941 Bateson and Mead had published “Principles of Morale Building”, which is the common root of Bateson’s “Morale and National Character” and Mead’s *And Keep Your Powder Dry. An Anthropologist Looks at America*, both published in 1942. In order to raise the morale, “an attitude to a group goal expressed in appropriate action”, they maintain, we must consider the different national characters, because the different cultural systems rely upon different sets of motives (unlike the German, the American morale increases with increasing adversity).<sup>84</sup>

Bateson’s and Mead’s problem is how the energetic commitment to the group purpose can be carried over from one situation and goal to another.<sup>85</sup>

---

Barker, “Preface” to *National Character and the Factors in its Formation*, 4<sup>th</sup> ed., London: Methuen, 1948 (1<sup>st</sup> ed. 1927), x).

81 “An account of the Select Society of Edinburgh”, *The Scots Magazine*, Vol. 17 (1755): 126b-127a.

82 Edinburgh NLS, Adv. Mss. 23.1.1, ff. 32–33, 35, 48, 51, 103, 190.

83 G. Bateson, *Morale and National Character*, in *Civilian Morale*, edited by G. Watson, Boston – New York: Reynal & Hitchcock, 1942, 71–91 (hereafter MNC).

84 G. Bateson, M. Mead, “Principles of Morale Building”, *Journal of Educational Sociology*, Vol. 15, No. 4 (1941): 206–220: 207–08, 211.

85 Ivi, 208.

Related to this problem of “morale transfer” is that of “morale resonance”; here their language evokes Humean “sympathy or contagion of manners”, a disposition that “makes us enter deeply into each other’s sentiments, and causes like passions and inclinations to run, as it were, by contagion, through the whole club or knot of companions” (E-NC, 202). “It is common – Bateson and Mead critically observe – to speak of emotional attitudes as ‘contagious’ or ‘infectious,’ but these terms are probably misleading”.<sup>86</sup> The bases for most of our adult attitudes derive from our family relationships “in childhood”; yet, when we speak of “contagious” morale attitudes we usually refer to the rapid spread of a behaviour pattern, like panic or determination, from one individual or group to another. The necessary basic attitudes were “already in some sense latent” in the individuals concerned, “ready to respond” when some other individual gave them overt expression.<sup>87</sup> To refer to these attitudes, Bateson and Mead maintain, we need “some concept of *latent morale*”,<sup>88</sup> since the rapid spread or evocation of latent attitudes is “a sort of resonance rather than a sort of contagion”.<sup>89</sup>

Education establishes a set of “coherent and socially adaptive” attitudes and values through a process that occurs “with greatest rapidity in early childhood” and contributes to the shaping of the character as a “conglomerate of attitudes”.<sup>90</sup> Each culture implants a set of attitudes and values in its individuals, but not all cultures implant the “*same*” set, and these attitudes are part of a coherent character structure that is adapted to the nation living.<sup>91</sup> Thus, to raise the morale we must consider the coherent attitudes which exist in a nation and are ready to be evoked. Knowing the national character we can predict how people will behave in wartime and adjust their interpersonal differences so that they can carry on war effectively.<sup>92</sup>

In order to achieve this result we must rely on the “habitual patterns” that people exhibit in their “continual dialectic relationships” of love and hate. The patterns show a “symmetrical” motif, where the individual is stimulated to respond with greater strength to the strength of the enemy; a “complementary” motif, where he is stimulated to respond with greater strength to the weakness of the enemy; and a “compensatory” motif, where he is stimulated to prove his uncertain strength by a symmetrical and complementary manner. As the symmetrical motif characterizes the Americans and the English, it is the most proper to build up their morale.<sup>93</sup> Contrary to the Germans, they are

---

86 Ivi, 208.

87 Ivi, 208.

88 Ivi, 209.

89 Ivi, 213.

90 Ivi, 213–214.

91 Ivi, 213–214 e n. 8.

92 Ivi, 216.

93 Ivi, 216–217.

positively stimulated by failure taken as a challenge, their morale will rise, and they react to it increasing the effort.<sup>94</sup>

As Bateson remembers in the *Ecology of Mind*, in 1942 “many of us were interested in ‘national character’”.<sup>95</sup> The practical consequences of the theory were evident; the abstract theory was entrusted to “Morale and National Character”.

### “The theoretical structure”: Bateson and national characters

“In casual conversation – Bateson observes – it is popularly assumed that national groups have characteristic differences”. The illustration is partly surprising: “the British are thought to be low in sense of humor and in foresight, but arrogant and possessed of bulldog tenacity”. The consequence is seemingly cautious: “If such differences exist, they are of tremendous importance” (MCN, 71). Bateson is reacting to an editorial article published in *Nature* and based on a speech of Morris Ginsberg (MCN, 71 n.).<sup>96</sup> Sceptical scientific thinking, the editorialist remarks, should be reconciled with the demand for positive contributions:

It is as easy to show confusions, contradictions and over-simplifications in the popular conceptions as it is difficult to find scientific proof of the real differences which, in a confused way, they may be registering. What then should be the role of a scientific treatment of the question?<sup>97</sup>

We should be able to show what, “at their most plausible”, popular conceptions would be and “in what directions scientific proof or disproof could most profitably be sought”.<sup>98</sup> Unlike Fyfe, Ginsberg thinks it “a mistake to dismiss the idea of national character as a mere illusion”. In order to identify differences in national character he looks at the institutions and their background, without overlooking “the pitfalls in the way of deducing national character from national institutions”.<sup>99</sup>

Ginsberg’s views on the English and Germans, the editorialist remarks, reflect “subjective impressions” rather than “scientific conclusions”, but interpret a body of permanent material “available for further inspection”.<sup>100</sup> Explaining national character is “even more hazardous” than to identify

---

94 Ivi, 217 e n.12.

95 G. Bateson, *Steps to an Ecology of Mind*, New York: Ballantine Books, 1972, 155.

96 Ginsberg quotes Hume’s essay, cf. *Reason and Unreason in Society*, London: Longmans, Green & Co., 1947, 165 and n. 1.

97 “National Character”, *Nature*, Vol. 148, No. 3741 (July 1941), 31–33: 31a.

98 Ivi, 31a.

99 Ivi, 31b-32a.

100 Ivi, 32b.

them, and we can only claim to “a degree of probability”.<sup>101</sup> According to Ginsberg “historical and social conditions play a much greater part than genetic factors”, and national character “is, for all its relative stability, capable of enormous changes”.<sup>102</sup> The editorialist remembers that the needs of warfare require to study “not only the relatively enduring character of a nation, but also its more transient moods”: the social scientist must contribute to a “full understanding” of the characters, moods and interests of the nations “among which the War is being fought, and by which an international order must be reconstituted”.<sup>103</sup>

Bateson acknowledges that his “abstract frame” for the research on national differences “will be useful only in so far as it suggests hypotheses which can be empirically tested” (MNC, 73). The American and English characters demand a different kind of propaganda “based on an understanding of their different ‘national’ psychologies” (MNC, 71). Bateson does not only show the differences in habit between various groups conditioned by cultural environment rather than by racial descent, but also their “implications [...] for the common character of the individuals who exhibit them” (MCN, 72). In order to state the “conceptual limits” within which the phrase is “likely to be valid”, he first examines some criticisms urged against “national character”; then, he outlines “what orders of difference we may expect to find” among western nations and guesses “concretely” at some of these differences; finally, he considers how the problems of morale are “affected” by them (MNC, 73–74).

Like Hume, Bateson allows the existence of a national or “common” character (E-NC, 203; MNC, 77, 79–80); yet, while Hume considers it as “uniform” (E-NC, 199, 204, 206), Bateson rejects any “uniformity” and limits himself to “regularity” (MNC, 75, 78, 89). The first part of Bateson’s argument recalls Hume’s opening distinction between national prejudices and characters. The examination of the “barriers” to national character “very stringently” limits its scope but also establishes its legitimate use (MCN, 79). Bateson accounts for two objections: the first appeals to the existing circumstances and historical background, which should be sufficient to account for the differences without invoking any unobservable difference of character (MNC, 74–75); the second appeals to differentiation, and articulates itself into five under-objections: three concern the differentiation within the nation or community (sexes, classes, and occupational groups; heterogeneity in melting-pot communities; accidental individual deviants); two concern the differentiation depending on changes (those within the nation and those of its boundaries).

The first objection, Bateson maintains, ignores “*learning*” and “learned character”: our characteristics “depend upon the previous experience and

101 Ivi, 32b.

102 Ivi, 33a.

103 Ivi, 33b.

behaviour” (MNC, 75). With regard to the second objection, some people argue that communities contain a “too great” internal differentiation or random element for any notion of common character; Bateson replies that his approach is useful, provided we describe this character in terms of “the motifs of relationship” between the differentiated sections, groups and individuals within the community (MNC, 77, 79), and provided we allow sufficient time for the community “to reach some degree of equilibrium or to accept either change or heterogeneity as a characteristic” (MNC, 79).<sup>104</sup> The postulate is that, like an individual, a nation is an “*organized*” entity, and all its aspects are “mutually modifiable and mutually interacting” (MNC, 76).

The insistence on the “mutual relation” within the community and between different communities, the “mutual relevance” between the special characteristics of different groups (MCN, 76), the “reaction” and “systematic relationship” among patterns (MNC, 78–79), recalls Hume’s insistence on “communication”:

where a number of men are united into one political body, the occasions of their intercourse must be so frequent, for defence, commerce, and government, that, together with the same speech or language, they must acquire a resemblance in their manners, and have a common or national character, as well as a personal one (E-NC, 202–203).

Hume considers communication both within a nation (E-NC, 204, 205, 209) and among different nations. The “situation of a nation with regard to its neighbours” (E-NC, 198) is one moral cause: “where several neighbouring nations have a very close communication together, either by policy, commerce, or travelling, they acquire a similitude of manners, proportioned to the communication” (E-NC, 206; cf. 215).

He who wants to describe the “common character of individuals” in a community must describe it in terms of “bipolar adjectives” and take “the dimensions of the differentiation as our clues to the national character”. Instead of defining the character of the Germans by its position on a continuum between extreme dominance and submissiveness (they are submissive), he will use some continua as “orientation towards dominance-submission” (they are dominant-submissive). By saying that they are “paranoidal” he will mean “some bipolar characteristic of German-German or German-foreign relationships” (MNC, 80, cf. 77).

Bateson displays his updated version of Humean sympathy: having discussed complementary (Germans) and competitive symmetrical (British and Americans) patterns, and combinations of complementary motifs (dominance-submission, exhibitionism-spectatorship, succorring-dependence), he happily concludes:

---

<sup>104</sup> According to Bateson “glorying in heterogeneity for its own sake” can be considered as a “common motif of behavior” (MNC, 77); according to Hume having “the least” of national character “may pass for such” (E-Nc, 207).

Using the motifs of interpersonal and intergroup relationship as our clues to national character, we have been able to indicate certain orders of regular difference which we may expect to find among the peoples who share our Western civilization (MNC, 89).

The knowledge of characters is useful as far as it tells us what “we may expect to find” (cf. MNC, 74–75, 78, 88–89). As Hume put it:

some particular qualities are more frequently to be met with among one people than among their neighbours [...]. We have reason to *expect* greater wit and gaiety in a Frenchman than in a Spaniard; though Cervantes was born in Spain (E-NC, 197–198; emphasis added).

## Conclusion

Nothing serves more to oppose any essentialist conception of nations and national characters than studying their theory and history. Especially because, even when we lose the feeling that our nation has a single identifiable character, we still ascribe it to others: we are “eager to give a national character”, as Wollstonecraft denounced. Hume had already warned us: “Men run with great Avidity to give their Evidence in favour of what flatters their Passions, and their national Prejudices” (Letters 2: 311). Godwin, Baroja, and Bateson: each of them makes a legitimate though partial use of Hume’s essay. Godwin and Baroja explicitly appeal to it. Bateson never mentions it, but his treatment is probably closer to Hume’s intention: to reject the prejudices without rejecting the characters, and to offer a plausible account of the latter in order to promote the good of the country.

As part of the vulgar we are inclined to form undistinguishing judgments and share national prejudices; as persons of good sense and reflection we correct these judgments and reject the prejudices, and allow the existence of (changeable) national characters and probable or frequent qualities and manners; finally, as prudent people we draw the practical consequences into our expectations, but we are equally ready to make exceptions; as Humeans we explain these characters by “moral” causes and “sympathy”, without forgetting that chance has an influence and only “the great liberty and independency, which every man enjoys, allows him to display the manners peculiar to him” (E-NC, 207).

CIP – Каталогизација у публикацији  
Народна библиотека Србије, Београд

1

FILOZOFSKI godišnjak = Belgrade philosophical  
annual / editor Slobodan Perović. – God. 1, br. 1 (1988)–  
– Belgrade : Institute of Philosophy, Faculty of Philosophy,  
1988- (Belgrade : Službeni glasnik). – 24 cm

Godišnje. – Glavni stvarni naslov od br. 28 (2015) Belgrade  
philosophical annual. – Tekst na engl. jeziku.

ISSN 0353-3891 = Filozofski godišnjak

COBISS.SR-ID 15073792

