

HUME – LECTURE NOTES & READING LIST

Lecturer: Dr. Norva Y.S. Lo

Hume Topics

Introduction to Hume

1. Reason vs. Passion
2. Moral Motivation – reason and passion
3. Moral Foundations – reason, passion, sympathy
4. Impressions and Ideas
5. Association of ideas
6. What Can We Know: two kinds of inquiry or reasoning
7. Sympathy & Moral Sentiments
8. Animals & Nature: from the Humean point of view

Selected Readings for Hume Topics

For each Hume topic, there is a set of readings **selected** from the following texts. Please see each Hume topic below.

 Baillie, J. 2000, *Hume on Morality*, London: Routledge.

 Hume, D. 1739-1740, *A Treatise of Human Nature*.
Freely available online at: <http://www.davidhume.org/texts/thn.html> .

 Hume, D. 1740, *Abstract of a Book lately Published - entitled A Treatise of Human Nature*.
Freely available online at: <http://www.davidhume.org/texts/abs.html> .

 Hume, D. 1748, *An Enquiry concerning Human Understanding*.
Freely available online at: <http://www.davidhume.org/texts/ehu.html> .

 Lo, N.Y.S. 2009, “Is Hume Inconsistent? – Motivation and Morals”, in C. Pigden (ed.), *Hume on Motivation and Virtue*, Palgrave Macmillan, pp. 57-79. Freely available on LMS subject website.

Hume Essay Questions (Answer one of the two questions below.)

Question 1. If you choose to answer this question, you must answer both part (a) and part (b), and given reasons for your answers.

- (a) Hume argues that there are only two kinds of inquiry or reasoning. What are they? What are the crucial differences between them?
- (b) Given Hume’s view on the two kinds of reasoning, could he consistently accept the proposition that people can do moral reasoning (e.g., reasoning about what is morally right and what is morally wrong)?
 - If you think that Hume can consistently accept the proposition, then please identify and explain under which one of the two kinds of reasoning Hume should classify moral reasoning?
 - If you think that Hume cannot consistently accept the proposition, then please explain why he cannot.

Question 2. If you choose to answer this question, you must answer both part (a) and part (b), and given reasons for your answers.

- (a) What does Hume mean by the terms “perceptions”, “impressions” and “ideas”? How do these three categories relate to each other? How, or via what capacities/faculties of ours, do we acquire impressions and ideas? Make up some examples of your own to illustrate your answers to the above questions.
- (b) A colour chat of blue wall paints is supposed to have six different shades of blue on it. The shades are placed next to each other from the lightest to the darkest shade. But due to printing error, the fourth shade of blue is missing (see below).



Suppose a person who has never previously seen (i.e., never had an impression of) the particular missing shade of blue is now looking at the chat. Does Hume’s theory of impressions and ideas *allow the possibility* that the person could work out in the mind (i.e., have an idea of) what the particular missing shade of blue would look like? Please explain your answer within the confines of Hume’s theory of impressions and ideas – whether you answer the question positively or negatively.

Introduction to Hume (Background Readings: Baillie 2000, chapter 1)

If Descartes is the father of modern philosophy, Hume is the person who gave shape to the contemporary philosophy world by first of all querying Descartes theories about knowledge, and then developing his own modest account of knowledge, and later his theories of ethics and aesthetics. During his lifetime, Hume was famous as a historian and intellectual, much loved in French academic circles, where he was known as “David the good” (*le bon David*). His sceptical critique of religion led to him failing to be appointed to a chair in philosophy in either Glasgow or Edinburgh, but he was close friends with Adam Smith who was appointed professor of logic at Glasgow in 1751. While Smith is well known today as contributing to the foundations of economics in his 1776 treatise *The Wealth of Nations*, his 1759 book *Theory of Moral Sentiments* was regarded as more significant by Smith himself. Hume appointed Smith as executor of his will.

Hume lived from 1711 to 1776, and left us a brief autobiography “My Own Life” (accessible through the subject website). Though born in Scotland, and becoming a student in Edinburgh at the young age of twelve, Hume subsequently moved to France so that he could live on a modest income. He settled in the little village of La Flèche, where both Descartes and Mersenne had, in the previous century, attended the famous Jesuit college. It was in La Flèche that he composed his most influential work, *A Treatise of Human Nature* (written between 1734 and 1737). The first volume (“of the Understanding”) and the second volume (“Of the Passions”) was published in 1739, and the third volume (“Of Morals”) was published in 1740. His best-known later summary works are his 1748 *Enquiry Concerning Human Understanding* (the best introduction to his overall system of philosophy) and his 1751 *Enquiry Concerning the Principles of Morals* (a work that Hume himself considered the best of all his writings). Hume also wrote copiously on history, and published many essays on moral, political and historical topics. By publishing a best-selling six-volume *History of England* over the years 1754 – 1762, Hume eventually became financially independent. Hume’s apparent atheism (or scepticism at the least), and his controversial views on topics such as suicide, ensured that he was regularly subject to attacks and machinations designed to prevent him getting posts of influence (hence his failure to secure any university academic position) and even aimed at getting him fired from posts he did occupy. As a librarian to the Faculty of Advocates at Edinburgh, for example, he became the target of attempts to dismiss him for having ordered “indecent books”, and a theologian who detested Hume’s liberal view on suicide threatened to take legal action against Hume’s publisher if the essay “On Suicide” were to be put in print. The essay was finally published after Hume’s death.

Many of those who knew Hume loved him, regarding him as a modest and friendly person who wore his impressive learning well. Towards the end of his life, he became friendly with a talented and mischievous young woman, Nancy Orde, whose father was an influential Scottish politician. One night, Nancy playfully chalked on the outside wall of Hume’s house in Edinburgh the words “St David’s Street”, the local nickname for the street in which *le bon David* lived. Subsequently the name of the street was indeed changed to “St David’s Street”, a gesture that would have amused Hume greatly. After Hume’s death, Adam Smith wrote: “... I have considered him both in his lifetime and since his death, as approaching as nearly the idea of the perfectly wise and virtuous man as perhaps the nature of human frailty will permit.”

In the 84 years from the death of Descartes to the birth of Hume, the modern science to which Descartes had devoted much of his study had increased in its power to the extent that the eighteenth century is now regarded as the *Age of Enlightenment* (or the Age of Reason) – an age when science promised to free people from false belief, ignorance and superstition and open up a new era of progress and discovery for humankind.

The Descartes lecture notes refer to the following tendencies as characteristic of early modern thought:

Mechanism: the universe, the bodies of animals, and the structure of plants all seen in terms of mechanisms – the comparison to clockwork was a common analogy.

Rationalism: the rules specifying how the world works are accessible to innate reason, and so the world is rationally intelligible. Appeals to magic and mystery were dismissed as irrational.

Scepticism: given that it was hard to distinguish among witchcraft, magic, religion and science, then the truth is not obvious and we have to be careful about what we claim to know.

Intellectual freedom: the combination of scepticism with the newly emerging sciences – especially physics and mechanics – led thoughtful people to challenge the authority of the Catholic church, Aristotle and the Scholastics philosophers. At the same time, thinkers embraced a new faith: in the power and liberating force of newly emerging sciences, especially physics

By the time of Hume, these tendencies had been accentuated and settled into core doctrines of Enlightenment thought, except for the second – rationalism, and the associated innatism.

The appeal to innate knowledge and ideas made by Descartes, was looking very shaky indeed by the time of Hume. The English philosopher John Locke in his 1690 *Essay concerning Human Understanding* delivered a scathing attack on the idea that there could be any innate knowledge of the kind Descartes imagined. He claims that innatists like Descartes believed that some ideas are innate in us (e.g., the idea of God, the idea of perfection, the idea of the infinite) because they thought that everyone had such ideas. In other words, if everyone agrees to a certain truths or ideas, the Cartesians would regard this as evidence that such truths and ideas are innate. But Locke had two objections to this. First, even the ideas of God and the infinite are not universal among people: “children and idiots,” he wrote. “have not the least thought of them”. Second, even if there were universal assent to some idea, this would not show the idea was born in us.

For Locke, there are rational and innate capacities in the human mind. He seems to have thought that Descartes confused the capacity to come to know something (say a theorem in geometry) with the knowledge of the theorem itself. We can have an innate *capacity* to learn geometry; but this does not mean that we were born with knowledge of Pythagoras’s Theorem. Stung by this attack on innatism, the German rationalist philosopher Gottfried Leibniz replied to Locke in his *New Essays on Human Understanding*, composed in 1704, but not published until nearly fifty years after Leibniz’s death. Leibniz’s argument was simple and harked back to an old argument put forward in classical times by Plato. Think of the way that we can arrive at definitions and proofs in geometry on the basis of drawing very irregular diagrams. Any hand-drawn triangle will likely have uneven sides, and its interior angles – if measured by precision tools – will certainly not add up to 180° . The truth of a theorem does not depend on our experiences of instances of badly-drawn triangles, so the proof of necessary truths in mathematics “can only come from inner principles”. Locke, by contrast, regarded the mind as an empty chamber or a blank slate, and what was in the mind could only come from experience. This is why Locke is called an “empiricist” philosopher – all knowledge coming from experience. Leibniz, like Descartes, is a “rationalist”, believing that capacities of the mind and knowledge of principles of reasoning are already in it at birth (innate knowledge) and that without the knowledge of innate principles we would have none of the universal, mathematical and theoretical knowledge that is needed in order to have a scientific understanding of the world. This debate between empiricists and rationalists continues, in one form or another, to the present day.

While a rationalist such as Descartes looks to explain certainty, and can find no basis for this in sensory experience, an empiricist like Hume regards *probability* rather than certainty as the “very guide to life” (the phrase comes from another 18th century philosopher, Joseph Butler, but clearly impressed Hume, who echoes it several times in his work). Yet, in a sly twist to the notion of innateness, Hume considers our *capacity* for having sense perceptions to be “innate” in the sense that sense perceptions are the *natural* happenings of the mind. Seeing a bright light, feeling a strong emotion, and other such “impressions” are experiences that naturally happen to us, and so our capacity for such experiences must be something born into us, although none of those particular experiences themselves are innate. Hume’s empiricism is an empiricism of concepts: for the concept empiricist, there can be no ideas that are not derived from, or in some way traceable back to, experience. While experience is a necessary ingredient in the formation of any knowledge, experience alone may not be sufficient to deliver certain kind of knowledge.

1. Reason vs. Passion (Readings: Hume T 2.3.3¹, T 3.1.1, M app 1.21²)

Hume has said some very provocative things about the roles of, and the relations between, reason and passions. For example:

“We speak not strictly and philosophically when we talk of the combat of passion and of reason. *Reason is, and ought only to be the slave of the passions*, and can never pretend to any other office than to serve and obey them.”³

“’Tis *not contrary to reason* to prefer the destruction of the whole world to the scratching of my finger.”⁴

What is Hume’s argument for these claims he made about reason and passions? In particular, what did he mean by “Reason is, and ought only to be the slave of the passions”? Here is David Lewis’s reading of Hume:

“In the first place, Hume’s ‘passions’ are sometimes none too passionate. He speaks of some passions as ‘calm’. We would do best to speak of all ‘passion’, calm and otherwise, as ‘desire’. In the second place, we call someone ‘reasonable’ in part because his desires are moderate and fair-minded. But when we do, I suppose we speak not strictly and philosophically. Strictly speaking, I take it that *reason* is the faculty in charge of *regulating belief*. And so I read Hume as if he had said that belief is the slave of desire. Our actions do, or they ought to, serve our desires according to our beliefs.”⁵

Hume’s view on the roles of reason and passions has had great influence on contemporary thinking in decision theories and moral philosophy. Lewis’s take on beliefs vs. desires is a contemporary application of Hume’s early ideas. While Hume distinguished different kinds of passions (e.g., love, regard, approval, or hatred, disgust, contempt), all of which play a role in motivating human behaviours, many theorists nowadays, like Lewis, use the general term “desire” to group together various feelings of attraction towards something, which are positive desires, or feelings of repulsion away from something, which are negative desires.

Following Hume, what Lewis is saying above is that the term “reasonable” is often used quite loosely or informally to describe a person who has moderate passions and desires. But strictly speaking, the term should be used to evaluative how well a person’s *rational* faculty “Reason” is functioning – not what the person’s desires are like. Strictly and philosophically, according to Hume, the term “reasonable” should be taken to mean “rational”, and a rational person is someone who reasons well – that is, someone who does valid or otherwise legitimate reasoning. Whether the person has moderate desires is quite a different and separate matter from whether the person is rational, or so would Hume and Lewis argue.

As Lewis points out, Hume takes reason to be the “faculty in charge of regulating belief” – i.e., it functions to work out which beliefs of ours are true or justified, and which ones are false or unjustified. So, we can think of a rational person as someone whose reasoning is effective with respect to acquiring true beliefs and getting rid of false beliefs. Whether such a person will behave morally is quite a separate issue, according to Hume. In short, a person may know the good but not desire the good, and even if a person desires the good, the desire might not be strong enough to result in the person actually doing the good.

Let us look deeper into the idea of being rational. What exactly is to be rational – in Hume’s view? What kinds of things can our rational faculty do? And what is its *limitation*? We can get some good ideas by looking at Hume’s account of reason in relation to human motivation and action. On the limitation on what reason can do, Hume says:

“I shall endeavour to prove *first*, that reason alone can never be a motive to any action of the will; and *secondly*, that it can never oppose passion in the direction of the will.”⁶

¹ The notation “T 2.3.3.” refers to Hume’s *A Treatise of Human Nature*, book 2, part 2, section 3. In general, “T n1.n2.n3.n4” would refer to book n1, part n2, section n3, paragraph n4 of the text. This notation system for the text is used through out the Hume lecture materials. The online text is also notated in this way.

² The notation “M app 1.21” refers to Hume’s *An Enquiry concerning the Principles of Morals*, appendix 1, paragraph 21. This notation system for the text is used through out the Hume lecture materials. The online Hume text are also notated in this way.

³ T 2.3.3.4, emphasis added.

⁴ T 2.3.3.6, emphasis added.

⁵ David Lewis, “Desire as Belief II”, *Mind* 105 (1996): 303, emphasis added.

⁶ T 2.3.3.1, emphases original.

On Hume's account of the human understanding, there are two, and only two, types of reasoning. The first is *deductive* reasoning, or what Hume calls "demonstrative" reasoning. This type of reasoning concerns conceptual relations, or what Hume calls "relations of ideas". For example, from the premise that David is at the library or in his office, and the second premise that David is not at the library, we can deductively infer that David is in his office. Any deductive inference from the premises "A or B" and "not A" to the conclusion "B" is valid – because of what the terms "or" and "not" means, and the conceptual relations between those terms, or what Hume calls "relations of ideas". The second type of reasoning recognized by Hume is *inductive* reasoning, or what he calls "probabilistic" reasoning. It concerns the states of affairs in the empirical world of sense experience, or what Hume calls "matters of fact". Predictions of election results based on newspaper polls, for example, or based on a certain percentage of vote counts after the election, is a form of inductive reasoning.

Now, in order to show that "reason alone can never be a motive to any action of the will", Hume needs to show that neither deductive reasoning nor inductive reasoning could by itself produce any motivation for action. At the core of Hume's argument is his premise that *passions* (or "desires" as Lewis puts it) are the *ultimate source* of motivation for action. Hume recognizes that reason can indeed help to move us to act in one way or another. But without the existence of some passion or some desire for a certain result, reason alone is never sufficient to motivate action. Deductive reasoning, Hume agrees, is useful in almost every art and profession. For example, given the information from various banks on their mortgage packages, I can by deductive mathematical reasoning work out which bank is offering the best mortgage deal overall. However, if I do not desire (or, to use Hume's term, do not have a "passion" for) some property, knowing where the best mortgage deal is would not motivate me to sign a mortgage contract with any bank. There has to be some existing passion in us for an object, Hume argues, in order for us to be motivated to take action about it. The passion is the source of motivation. Or, we may say, the passion *is* the motivation. Hence, deductive reasoning concerning relations between concepts (e.g., mathematical relations) can *never by itself* have any influence on our actions. It can direct us to choose an effective means or action to achieve some end or purpose – but only if we already desire that end or purpose.⁷

What about inductive reasoning concerning matters of fact, such as causes and effects? Can inductive reasoning by itself produce any motivation for action? Hume's answer is negative too, and for the same reason – namely, passion is the only source of motivation for action.

"'Tis obvious, that when we have the prospect of pain or pleasure from any object, we feel a consequent emotion of aversion or propensity, and are carry'd to avoid or embrace what will give us this uneasiness or satisfaction. 'Tis also obvious, that this emotion rests not here, but making us cast our view on every side, comprehends whatever objects are connected with its original one by the relation of cause and effect. Here then [inductive/probable] reasoning takes place to discover this relation; and according as our reasoning varies, our actions receive a subsequent variation. But 'tis evident in this case, that the impulse arises not from reason, but is only directed by it. 'Tis from the prospect of pain or pleasure that the aversion or propensity arises towards any object: And these emotions extend themselves to the causes and effects of that object, as they are pointed out to us by reason and experience. It can never in the least concern us to know, that such objects are causes, and such others effects, if both the causes and effects be indifferent to us. Where the objects themselves do not affect us, their connexion can never give them any influence."⁸

For example, due to past experience, a teacher may by inductive reasoning work out the causal relation between giving high marks to student assignments and receiving high teaching survey scores from the students in return. But if the teacher has no desire at all for receiving high survey scores, or if he has such a desire but it does not override his other desires that pull him to a different direction (e.g., a desire to give marks in accordance with quality, or a desire to be someone with integrity), his believing in the causal relation between the marks he gives and the survey scores he receives, this belief by itself, would have no influence on the ways he marks student assignments.

In short, reason is the faculty in charge of producing and evaluating beliefs – via deductive or inductive inferences, or a combination of both. From this, Hume concludes the following:

⁷ See T 2.3.3.2.

⁸ T 2.3.3.3

[R]eason, in a strict and philosophical sense, can have an influence on our conduct only after two ways: Either when it excites a passion by informing us of the existence of something which is a proper object of it [i.e., the passion]; or when it discovers the connexion of causes and effects, so as to afford us means of exerting any passion.⁹

“[’T]is only in two senses, that any affection can be call’d unreasonable. First, when a passion, such as hope or fear, grief or joy, despair or security, is founded on the supposition of the existence of objects, which really do not exist. Secondly, when in exerting any passion in action, we chuse means insufficient for the design’d end, and deceive ourselves in our judgement of causes and effects. Where a passion is neither founded on false suppositions, nor chuses means insufficient for the end, the understanding [i.e., reason] can neither justify nor condemn it. [...] In short, a passion must be accompany’d with some false judgement, in order to its being unreasonable; and even then *’tis not the passion, properly speaking, which is unreasonable, but the judgement.*”¹⁰

That is why Hume says “’Tis not contrary to reason to *prefer* the destruction of the whole world to the scratching of my finger” – provided that, we may add, such a preference (which is a passion) is not founded on, or accompanied by, the supposition of some “false judgement”. But even if it is, it is not the preference, “properly speaking, which is unreasonable, but the judgement”.

Eleven years after the third Book of his *Treatise* (1740), Hume’s *Enquiries concerning the Principles of Morals* was published – a work that he himself considered to be the best of his. There, Hume maintains the same distinction between reason and passions:

“The distinct boundaries and offices of reason and of taste are easily ascertained. The former conveys the knowledge of truth and falsehood: The latter gives the sentiment of beauty and deformity, vice and virtue. The one discovers objects as they really stand in nature, without addition or diminution: The other has a productive faculty, and gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises, in a manner, a new creation. Reason, being cool and disengaged, is no motive to action, and directs only the impulse received from appetite or inclination, by showing us the means of attaining happiness or avoiding misery: Taste, as it gives pleasure or pain, and thereby constitutes happiness or misery, becomes a motive to action, and is the first spring or impulse to desire and volition.”¹¹

⁹ T 3.1.1.12.

¹⁰ T 2.3.3.6, emphases added.

¹¹ M app 1.21, where “taste” is the faculty in charge of the passions.

2. Moral Motivation: reason and passion (Readings: Lo 2009, Hume T 3.1.1-2, T 3.2.1-3, T 3.3.1, M app 1.21)

According to Hume, reason *alone* can never determine the distinction between moral good and evil. That is to say, we can *never* find out whether an act is morally right or wrong *just* by using our reasoning - whether it is deductive reasoning, or inductive reasoning, or a combination of both. Among Hume's many arguments for this anti-rationalist thesis on the nature of morality and moral inquiry is what is often nowadays called the "motivation argument". The passage below gives some background to the pre-dominant rationalist tradition, which Hume's motivation argument is designed to attack.

"Those who affirm that virtue is nothing but a conformity to reason; that there are eternal fitnesses and unfitnesses of things, which are the same to every rational being that considers them; that the immutable measures of right and wrong impose an obligation, not only on human creatures, but also on the Deity himself: All these systems concur in the opinion, that morality, like [demonstratively discoverable] truth, is discern'd merely by ideas, and by their juxta-position and comparison. In order, therefore, to judge of these systems, we need only consider, whether it be possible, from reason alone, to distinguish betwixt moral good and evil, or whether there must concur some other principles to enable us to make that distinction".¹²

In the passage above, moral inquiry, i.e., inquiry about moral good and evil, right and wrong, is compared to conceptual inquiry attainable by deductive reasoning about conceptual relations – or what Hume calls "demonstrative reasoning" about "relations of ideas". It says that there is a view out there suggesting that just by analyzing and understanding what the terms in a moral claim mean, our rational faculty could somehow work out whether the moral claim is true or false. For example, perhaps just by understanding what the phrase "telling a lie" means, and what the term "morally wrong" means, we can use reason to determine the truth or falsity of the moral claim "it is wrong to tell lies". It is as if moral claims are analytic claims – i.e., claims the truth or falsity of which can be determined just by analyzing the meanings of the terms they contain.

It is important to note that in the first passage above, Hume is not actually describing his own position. Rather he is characterizing a form of moral rationalism, which he questions, opposes, and later rejects by his motivation argument. That is why, at the end of the passage he says: "In order, therefore, to judge of these systems [i.e., the moral systems he questions], we need only consider, whether it be possible, from reason alone, to distinguish betwixt moral good and evil, or whether there must concur some other principles to enable us to make that distinction."

Given Hume's characterization, the kind of moral rationalism that he is opposing maintains the following four core claims:

- R1. Something is morally right if and only if it is rational, whereas something is morally wrong if and only if it is irrational.
- R2. Moral judgements, which distinguish between good and evil, right and wrong, are produced by reason alone. (from R1)
- R3. Humans are rational beings, who will develop passions for/against, and be motivated to do/avoid whatever Reason judges to rational/irrational.
- R4. Moral judgements, which distinguish between good and evil, right and wrong, will generate corresponding passions and motivate corresponding actions. (from R1, R2 and R3)

Proposition R1 captures what Hume means when he describes the moral rationalist as maintaining "virtue is nothing but a conformity to reason". For the rationalist, to be morally good is nothing but to be rational – in the sense of "conforming" to, or following, reason. On the other side of the same coin, moral evil is nothing but a form of irrationality, or violation of, contradiction to, reason.

We might wonder how exactly we are going to work out whether something conforms to or contradicts reason. The rationalist would reply that reason knows its own standard the best, and so it is by using our reason, and reason alone, that we can work out whether something (e.g., an action) is rational or irrational. But what is morally right is equivalent to what is rational, and what is morally wrong is equivalent to what is irrational – according to R1. So, it

¹² T 3.1.1.4.

follows that it is by using our reason, and reason alone, that we can work out whether something is morally good or evil. In short, R2 is derived from R1.

Next, under the picture of moral rationalism, a rational being is someone who not only knows what is rational to do but also desires, and therefore is motivated to pursue, what is rational. The third core claim of moral rationalism, R3, says that that human beings, like you and I, are all “rational beings”, whose emotions and behaviours are governed by their reason. In fact, the rationalist takes rationality to be the essence of humanity, that which separates human beings from, and indeed puts them above, nonhuman animals. So, we are human only to the extent that we are rational.

Finally, the last core claim of moral rationalism, R4, follows directly from the previous three moral rationalist claims. It says that moral judgements excite our passions and motivate our actions. But why will we be motivated by our moral judgements? According to the rationalist, that is because we are rational beings who are motivated by our reason’s judgements about rationality, and judgements about morality are nothing but judgement about rationality.

In fact, under the moral rationalist picture, it is not only that all “human creatures” are rational beings motivated by reason. God “the Deity himself” is also a rational being, and indeed a perfectly rational being. For the moral rationalist, the source of human immorality is therefore the inferior rational capacities human beings have. There are two ways for us to fail morally. Firstly, we may reason imperfectly, e.g., by mistaking an irrational course of action as a rational one. When we do that, we are, according to the rationalist, in effect mistaking the morally evil with the morally good. Secondly, we may fail morally when our reason fails to combat our passions when it fails to provide a motivation strong enough to result in actual moral behaviour, e.g., if we have an irrational desire to act otherwise, which is stronger than the rational motivation generated by reason.

For example, my reason might tell me that having another bar of chocolate is will put my weight up, providing me a motivation to avoid another bar of chocolate. But my desire for more chocolate might be a very strong one, and so override my rational judgement against more chocolate. The rationalist would say, in this case, my reason, being imperfect, has failed to combat my desire.

Hume, of course, would reject the rationalist way of describing the situation. Hume would argue that having a desire for something is not unlike feeling an itch, that feelings, passions or desires are not capable of being true or false, and so they are not the appropriate kind of objects to be judged by reason. Regarding the story about chocolate, what Hume would say is that there is my desire for more chocolate, and there is also my desire to maintain a certain weight. It is not reason that has failed when I finally take another bar of chocolate – but my desire to maintain a certain weight, or be health, has failed. It is because my desire for maintaining a certain weight, or be health, is calmer and weaker but my desire for chocolate is violent and strong that I end up eating more chocolate. For Hume, there is no such thing as combat between reason and desire – but only the combat between one desire and another desire. As we have seen in the previous lecture, Hume argues that in general, reason alone, without the presence of any passion or desire, cannot motivate action. In the chocolate story in particular, he would simply point out that if I do not care about my weight, or health, at all, i.e., if I have no desire to maintain a certain weight, or to be healthy, then no amount of reasoning is going to help motivate me to avoid eating more chocolate. This confirms Hume’s general thesis that reason alone cannot influence passions or motivate actions.

Let us now look at the original text from Hume where he puts forwards his motivation argument against moral rationalism. Hume writes:

“[M]orality is [...] *supposed* [e.g., by the moral rationalists] to influence our passions and actions [...] And this is confirm'd by common experience, which informs us, that men are often govern'd by their duties, and are deter'd from some actions by the opinion of injustice, and impell'd to others by that of obligation. Since morals, therefore, have an influence on the actions and affections, it follows, that they cannot be deriv'd from reason [alone]; [...] because reason alone [...] can never have any such influence. Morals excite passions, and produce or prevent actions. Reason of itself is utterly impotent in this particular. The rules of morality, therefore, are not conclusions of our reason. [...] An active principle can never be founded on an inactive; and if reason be inactive in itself, it must remain so in all its shapes and appearances”.¹³

¹³ T 3.1.1.5-7, emphasis added.

There are many variant interpretations of Hume's motivation argument. One popular interpretation of the argument can be summarized (in standard form) as follows.¹⁴

- H1. Moral judgements, which distinguish between good and evil, right and wrong, excite passions and motivate actions. (E.g., moral judgement of the form "action x is morally right (or wrong)" excites the sentiment of approval (or disapproval) in us towards action x, which then motivates us to (or not to) take action x).
- H2. Reason alone cannot excite passions or motivate actions.
- H3. That which excites passions and actions cannot be produced solely by a faculty which alone cannot excite passions and actions.
- H4. Moral judgements, which distinguish between good and evil, right and wrong, cannot be produced by reason alone. (from H1, H2 and H3)

Consider H1. This premise is actually one of moral rationalism's core claims, namely R4. This makes good sense of Hume's statement that "[M]orality is [...] supposed to influence our passions and actions". Supposed by whom? Suppose by the moral rationalist.

Consider H2. It states Hume's general thesis that Reason alone can never excite passions or motivate actions. This is a claim on the limitation and boundary of reason.

Consider H3. Hume seems to think that H3 is so self-evidently true that he (as far as I know) has never defended it. That is perhaps because he takes H3 to be a true "relation of ideas" – i.e., an analytic statement which is true simply by virtue of the meanings of the words it contains.

From the three premises H1, H2, H3, Hume arrives at the conclusion, H4, which says that moral judgements cannot be produced by Reason alone. That is to say, reason by itself can never determine the distinction between moral good and evil. We can never find out whether an act is morally right or wrong just by using our reasoning - whether it is deductive reasoning, or inductive reasoning, or a combination of both.

Hume's motivation argument is valid – i.e., under the supposition of his three premises H1, H2 and H3, the truth of his conclusion H4 logically follows. So, whether Hume's motivation argument against moral rationalism is successful will depend whether Hume's premises in the argument are true.

Let us now consider Hume's whole motivation argument, and compare it side by side with the core claims of moral rationalism. As we have seen, H1, the first premise in Hume's motivation argument, is actually identical to the rationalist claim R4. We can also see that H4, the conclusion in Hume's argument, is the negation (i.e., the direct opposite) of the rationalist claim R2. In short, **H1 = R4; H4 = rejection of R2**. So, what is going on?

What is going on in Hume's motivation argument is that he is actually employing the rationalist's own claim R4 as a premise, to argue against the rationalist's other claim R2. What Hume has in effect shown is that given his own premises H2 and H3, the rationalist thesis R4 implies the rejection of the rationalist thesis R2. In other words, if Hume's premises H2 and H3 are true, then moral rationalism is an inconsistent position in that not all of its core claims can be true at the same time.

Hume's motivation argument, which concludes with the rejection of R2, also works against another rationalist claim, namely R1, which says that the morally good is equivalent to the rational, and the morally evil is equivalent to the irrational. Why is the rejection of R2 also a rejection of R1? That is because R2, as we have seen earlier, deductively follows from R1. So if R2 is false or unacceptable, then so is R1.

In short, whether Hume's motivation argument against moral rationalism is successful or not crucially depends on whether his core premise H2 is true. As we have seen in the previous lecture, Hume argues that our own experience tells us that the existence of some passion in us is always a necessary condition for motivation and action. Some passions, according to Hume, are more fundamental than others in that we acquire them very naturally through the operation of the non-rational faculty which he calls the "taste". This faculty, "taste", in Hume's view, is constituted by

¹⁴ For more interpretations of Hume's motivation argument, see Lo 2009.

some *non*-rational (but not irrational) psychological principles, for example, the principles that we are naturally attracted to pleasure and objects that we believe to be the causes of pleasure, and that we are naturally repelled by pain and objects that we believe to be the causes of pain. These are principles which Hume believes to govern not only human passions, but also the passions in nonhuman animals. By contrast, reason's function (and its only function), on Hume's view, is to regulate our beliefs (i.e., to estimate their likelihood of being true or false). In particular, reason alone can never produce any original passions in us, and at best can only re-direct some already existing passion into choosing a more effective means to satisfy itself. For example: someone may have a desire for money and fame because they believe that money and fame are the means that will bring them love and friendship, which are the ultimate ends they fundamentally desire. Whether the belief that money and fame will bring love and friendship is true or false is to be discovered by reason – e.g., inductive reasoning. But what original or most fundamental passions or desires people have is governed by taste, not reason.

If reason knows its standard the best as the rational thinks, then reason should know its own limitation even better. According to Hume, who is a very rational person – although far from being a rationalist, H2 states the limitation of reason, namely: Reason alone can never excite passions or motivate actions. That is why he says that “reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them”. Here I read Hume as if he is saying: If we are truly rational, then we will, and we *rationally ought* to, know the limitation of reason.

Now, if Hume is right in arguing that reason alone is not sufficient to produce moral judgements that distinguish between good and evil, right and wrong, then what is the missing ingredients? This question takes us to Hume's positive account on the nature and foundations of human morality, which we will look at in the next lecture.

3. Moral Foundations: reason, passion, sympathy (Readings: Lo 2009, Hume T 3.1.1-2, T 3.2.1-3, T 3.3.1, M app 1.21)

Hume's motivation argument against moral rationalism concludes that moral judgements, which distinguish between good and evil, which motivate human behaviours, cannot be the products of Reason alone. If Hume is right in arguing that our rational faculty by itself is unable to produce moral judgements, then what is missing from the moral rationalist picture? What other human faculties need to be involved in order for us to be able to make the distinctions between good and evil, to determine whether something is right or wrong? This question brings us to Hume's positive account on the foundations of human morality.

Hume account of the foundations of morality is given in the Book 3 of his *Treatise*, which is repeated in a simplified way in his later work, the *Enquiry concerning the Principles of Morals*. Hume's views and arguments on morals are based on, and further develop, his account of the human mind given in Book 1 and Book 2 of the *Treatise* – which includes his theory of the human *understanding*, and his theory of the human *passions*, and, in particular, his theory of *sympathy* as a higher level mental mechanism relying on the combined operations of various more basic principles that he believes to govern the human mind. (NOTE: Hume's idea of what he calls "sympathy" is equivalent to what we nowadays usually call "empathy".)

Methodologically, in order to find a way to decide whether a moral judgement is true or false, a Humean would first of all ask whether such a judgement is an analytic proposition concerning "relations of ideas" or an empirical proposition concerning "matters of fact". That is because Hume, remember, recognizes only these two kinds of propositions – there is not a third kind for him. How would Hume answer this question about the nature of moral judgements – are they analytic or are they empirical? What can find some clues by hunting through Hume's writings. Consider the following passages, for example, where Hume can be seen as giving a conceptual analysis, or definition, for the moral terms "vice" and "virtue":

"[W]hen you pronounce any action or character to be vicious, you *mean* nothing, but that from the constitution of your nature you have a feeling or sentiment of blame from the contemplation of it."^{15, 16}

"The hypothesis which we embrace is plain. It maintains, that morality is determined by sentiment. It *defines* virtue to be *whatever mental action or quality gives to a spectator the pleasing sentiment of approbation; and vice the contrary.*"¹⁷

Accordingly, Hume's conceptual analysis, or definition, of "virtue" and "vice" can be put approximately as this:

An action or character is virtuous/vicious if and only if a "spectator" feels a sentiment of approbation/disapprobation towards it.

Now, it is very important to notice that what Hume calls a "spectator" or "judicious spectator"¹⁸ is not just anyone. Rather, Hume's "spectator" is human being who – on top of being mentally normal (i.e., having a mind where the various basic principles of human nature as identified by Hume are functioning, including those underlie the operation of sympathy) – must also meet the following extra conditions:¹⁹

- C1.** Having full information of all the circumstances relevant to the case.
- C2.** Overlooking one's personal interests – if any is involved in the case.
- C3.** Overlooking one's personal relations – if any is involved in the case.
- C4.** Taking into account of facts about human nature.

¹⁵ T. 3.1.1.26, emphasis added.

¹⁶ According to Hume (T 3.1.2.2-6, T 3.1.2.11), the sentiment of approbation is a "pleasing sentiment", and disapprobation the contrary. Furthermore, Hume (T 1.3.10.2, T 2.1.1.4, T 2.1.10.8, T 2.3.3.3, T 2.3.9.1, T 3.3.1.2, and M app 1.21) maintains that we naturally desire what pleases us and are repelled by what brings us pain. It follows that the moral sentiments can excite desires and thus to some extent motivate moral behaviour. Cf. M app 1.3.

¹⁷ M app 1.10, first emphasis added, the rest original.

¹⁸ T 3.3.1.14.

¹⁹ See M app 1.11–3, T 3.1.2.3–4, T 3.3.1.15–8.

If you have a normal functioning human mind, and if you also satisfy all the above four extra conditions when considering a matter of morals, then you are qualified as a Humean “spectator” – i.e., someone who is in a good position to make reliable moral judgements about good and evil, right and wrong, virtue and vice.

In his pursuit of a “science” of human nature, which is the background to his account of human morality, Hume, like many of his contemporaries, believes that all members of the human species share the same nature, in that the mind of every human person is governed by the same set of basic principles (e.g., those that he describes and argues for in Book 1 and Book 2 of the *Treatise*). So, Hume expects that every human person with a functioning mind would feel the same and come to the same judgement about a matter of morals – should they consider the matter when they are under the same four conditions C1, C2, C3, and C4. Accordingly, Hume’s conceptual analysis or definition of “virtue” and “vice” can be expounded as follows:

H. An action or character is virtuous/vicious if and only if all human beings (with a normal functioning mind) are *disposed*, under the conditions C1, C2, C3 and C4, to *feel the sentiment* of approbation/disapprobation towards it.²⁰

According to Hume, it is via the operation of *sympathy* that we develop the sentiments of approbation and disapprobation towards social virtues and vices (e.g., justice and injustice), respectively.

Sympathy, as we will see later (Topic 7), in Hume’s view is a natural psychological mechanism of the human mind, by which the mind’s thoughts of the sentiments of others (e.g., my idea of another person’s pleasure) can be converted into its own sentiments of the same kind (e.g., my impression, i.e., my own feeling, of pleasure). For example, when I observe a smile on your face and laughter in your voice, I detect a joyful feeling in you. In response to that I can actually, by empathy, also feel a joyful pleasure in myself. When someone smiles to us, we often naturally respond with a smile back. When we hear other people laugh we often feel more uplifted and often laugh too. Likewise, when we detect the feeling of distress or fear in others (as in a melodrama or in a horror movie), we often naturally also develop a feeling of distress or fear in ourselves. These different emotional reactions are all results of our empathy in operation.

On the relevance of sympathy to moral judgements, Hume writes:

“[W]e never fail to observe the prejudice we receive, either mediately or immediately, from the injustice of others [... Even] when the injustice is so distant from us, as no way to affect our interest, it still displeases us; [... That is because] we **partake** of [the feeling of] uneasiness [or suffering in the victims of injustice] **by sympathy**; [...] “a **sympathy with public interest** is the source of the moral approbation. [... Sympathy] has sufficient force to influence our taste, and give us the sentiments of approbation or blame.”²¹

For Hume, the mechanism of empathy produces moral sentiments in us in response to the emotions we detect in other people - just as for a medical scientist, certain physiological mechanisms produce our experience of various bodily sensations (e.g., the sensations of colour, smell, sound, and heat) in response to external stimuli.

Now, **given H** (Hume’s conceptual analysis of morals), what do I need to do in order to make a correct moral judgement, for example, a judgement concerning whether one particular individual’s killing of another particular individual is a vicious act? The pre-condition is of course that I have a normal functioning human mind. If I satisfy the pre-condition, then I will need to proceed to putting myself into the four extra conditions C1, C2, C3, and C4 in the following manner:

Step (1) I get myself into condition C1, i.e., by *observation* and *reasoning*, I gather as much *information* as I can about the circumstances under which the act was committed and the relations between the actor and all the other people involved, for example, what the motive of the act was, whether it was an accident, a case of self-defence, or unprovoked, and also what histories and relationships the killer and the victim had, ... etc.

²⁰ According to Hume (T 3.2.1.2-8), an action is virtuous/vicious only if it is the product of a virtuous/vicious “motive or principle”. Hence, in the case of evaluating an action, condition C1 requires that the spectator is aware of the motive or principle, from which the action was produced.

²¹ T 3.2.2.24

- Step (2) I get myself into condition C2, i.e., I try as hard as I can to overlook my self-interests if any is involved in the case under consideration. For example, suppose the victim is my benefactor so that I suffer a material loss due to her death. I will pretend or imagine as hard as possible that my interests are not at stake.
- Step (3) I get myself into condition C3, i.e., I try as hard as I can to overlook my personal relations if any is involved in the case. For example, suppose the killer is a friend of mine. I will pretend or imagine as hard as possible that the killer were a stranger instead.
- Step (4) I get myself into condition C4, i.e., I take into account of common facts of human nature. For example, suppose the killer has a history of suffering from long term physical and psychological abuses, I pay attention to the fact that it is part of the human condition that people sometimes lose sanity or sensitivity after repeated traumatic events.

Now, after having successfully got myself into all the above conditions, I then quickly proceed to:

- Step (5) While I am still under those four conditions, C1, C2, C3 and C4, I inspect my passions to find out *what sentiment I actually feel* towards the act under evaluation.
- Step (6) Finally, I make a judgement about the moral quality of the act according to the sentiment I felt towards the act in step (5) above, and also according to H, Hume's conceptual analysis of morals. For example, if I felt the sentiment of *disapprobation* towards the act under the four conditions C1, C2, C3 and C4, then, given H, the act is (and I should make judge it to be) *vicious*. But if I did not feel the sentiment of disapprobation, then the act is not (and I should not judge it to be) vicious. Alternatively, if I felt the sentiment of *approbation* instead (e.g., because in step (1) I found good evidence showing that the act was taken against a mass killer who was in the middle of shooting innocent people at a school), then the act is (and I should judge it to be) *virtuous*.

The four conditions identified by Hume, C1, C2, C3 and C4, can be thought of as conditions favourable or *ideal* for developing reliable moral sentiments – in the sense that the better we get ourselves into those conditions, the better indicators of genuine moral properties our subsequent moral sentiments will be. If we could get ourselves completely into those conditions, then, Given H, our subsequent moral sentiments would be infallible indicators: our positive sentiments of approbation, delight and esteem would indicate virtue; whereas our negative sentiments would indicate vice. It is unlikely of course that we will completely satisfy those four conditions that are ideal for having reliable sentiments. So we approximate the test and get less than ideal but approximate results. We also try to improve the test by trying harder to meet those ideal conditions. This is how we make progress in our inquiry about morals. Hume's meta-ethical account of morals, like many other dispositional theories of values nowadays, "makes a place for truth, and in principle for certain knowledge, and in practice for less-than-certain knowledge about value. But also it makes a place for ignorance and error, for hesitant opinion and modesty, for trying to learn more and hoping to succeed."²² In short, under Hume's conceptual analysis of morals, H, the method (from step (1) to (6)) for moral inquiry is not mysterious, but *empirical* and potentially *scientific*.²³

It should be noticed that Hume's conceptual analysis of morals, H, is a *humanist* as well as a *universalist* analysis of morality. It is humanist in that it understands value and disvalue as essentially anchored on basic psychological dispositions of the *human being*, not that of God or any other superhuman beings. It is universalist in that it implies that something is valuable/disvaluable only if *all* human beings are disposed under the ideal conditions to feel a positive/negative sentiment towards it. On his universal humanism, Hume writes:

The notion of morals implies some sentiment common to *all mankind*, which recommends the same object to general approbation, and makes every man, or most men, agree in the same opinion or decision concerning it.²⁴

²² David Lewis, "Dispositional Theories of Values", *The Proceedings of the Aristotelian Society*, Supplementary Volume 63 (1989), p. 123.

²³ See M app 1.10, where after putting forward his conceptual analysis of morals, he says: "We then proceed to examine a plain matter of fact, to wit, what actions have this influence [on the sentiments of the judicious "spectator"]: We endeavour to extract some general observations with regard to these sentiments" (emphasis added). Also see *ibid.* 1.10.

²⁴ M 9.5 (emphasis added). The notation "M 9.5" refers to *An Enquiry concerning the Principles of Morals*, section 9, paragraph 5. This notation system for the text is used through out the Hume lecture materials. The online text is also notated in this way.

When a man [...] bestows on any man the epithets of *vicious* or *odious* or *depraved*, he then [...] expresses sentiments, in which, he expects *all* his audience are to concur with him.²⁵

[T]he sentiments, which arise from humanity, are not only the same in all human creatures, and produce the same approbation or censure.²⁶

It is perhaps his optimism in the uniformity of human nature (which is, after all, an empirical matter) that has led Hume to consider universality as a conceptual component of moral terms. Those Humeans who lack such an optimism (which is common for enlightenment thinkers) might want to hold a neutralized Humean (but not Hume's own) analysis of morals, which says:

H*. Something is a relative/universal value (or disvalue) if and only if some/all human beings with a normal functioning mind are disposed, under C1, C2, C3 and C4, to feel a positive (or negative) sentiment towards it.

²⁵ M 9.6 (last emphasis added, the rest original).

²⁶ M 9.7.

4. Impressions and Ideas (Readings: Hume E 2²⁷ and T 1.1.1 – T 1.1.3)

Hume divides all “perceptions” (i.e., experiences) into “impressions” and “ideas”. This theory device gives him a more finely grained account of the operations of the mind than either Locke or Descartes have. Impressions are *original* “perceptions of the human mind” which are vivid, forceful, strong or lively. Ideas are the “faint images” of the original impressions. In other word, ideas are derived or secondary. Hume’s “**copy principle**” states:

“A]ll our simple ideas in their first appearance are deriv’d from simple impressions, which are correspondent to them, and which they exactly represent” (T 1.1.1.7).

While it is natural for us to have impressions and ideas, Hume thinks the copy principle captures what is important in the empiricist claim that there are no innate ideas. Notice that copies can sometimes be more vivid than the original impressions, as when a photograph is enhanced on the computer so as to brighten up the original. So the point about impressions being more lively than ideas is one that holds generally – but not universally. More frequently, Hume often refers to ideas as representing impressions – a generally more helpful way to think about the distinction between the two categories: we attain impressions first, which then give rise to the corresponding ideas.

But the copy principle has to be treated with caution. It is hardly plausible, the rationalists would say, that all our ideas are simple copies of impressions. From what impression does the idea of the number 3, or the number 79, derive. And what about Descartes’s examples – ideas like those of God, or of a triangle, or of infinity? Hume sometimes writes as if the difference between impressions and ideas is the difference between directly experiencing something, on the one hand, and merely thinking about it on the other. If this is right, then having an impression of how a lemon tastes, for example, will involve actually tasting a lemon. But if I merely think about the taste of a lemon, without actually tasting one, then I will just have an idea, not an impression, of the taste of a lemon. Likewise, to have an impression of embarrassment, or of dizziness, is different from thinking about embarrassment or dizziness. But this analogy is limited, and it is unlikely that we can capture what Hume intends by the distinction between impressions and ideas by assimilating it to the categories of directly experiencing something versus thinking about it. Better, on the whole to remember that we have impressions and we have ideas:

- (a) all ideas are derived from, and so are traceable back to, some impressions,
- (b) ideas represent, by having contents resembling to, the impressions from which they are derived, and
- (c) generally but not necessarily, ideas are fainter or less vivid than the impressions from which they are derived.

Now, suppose I am given a piece of paper with an equilateral triangle on it. As I study the triangle, I see that it has three equal sides. Furthermore, I observe that it has three equal angles. When I stare at this particular equilateral triangle, I have an impression of it. When I close my eyes and recall what it I have observed, I have an idea of it. The same process occurs every time when I observe a new triangle. After a period of observing and studying different triangles (e.g., a scalene one, an isosceles one, or a right-angle one), I will have developed in my mind a collection of separate ideas of those different triangles.

An interesting question is whether from these individual ideas of different triangles, could I develop some so-called “abstract” or “general” idea of the triangle. According to Locke, “ideas become general by separating from them the circumstances of Time and Place ... By this way of abstraction they are made capable of representing more individuals than one; each of which, having in it a conformity to that abstract idea ...” (*Essay* III, iii, 6). In Locke’s view, in forming an abstract idea, one has to use the full powers of the mind to generate an idea of a triangle which is neither equilateral, nor isosceles, nor scalene – “but all and none of these at once” (*Essay* IV, vii, 9).

For Hume, Locke’s theory of abstraction is nonsense. Hume regards all ideas as copies of, or traceable back to, some impressions of particular things. We have many separate ideas of many different triangles, which are derived from a series of impressions we have had of them. Hume argues that if we try to think of some feature of triangles in general, a strange phenomenon would occur. Suppose we wonder whether all the angles of a triangle are always equal to each other (T 1.1.7.8), then immediately other ideas – e.g., an idea of some scalene triangle and then an idea

²⁷ The notation “E 2” refers to Hume’s *An Enquiry concerning Human Understanding*, the entire section 2. The notation “E 2.1” would refer to section 2, paragraph 1. This notation system for the text is used through out the Hume lecture materials. The online text is also notated in this way.

of some isosceles triangle – “immediately crowd in upon us, and make us perceive the falsehood of this proposition”. So, instead of saying that we have an abstract idea of the triangle which has all the properties of all triangles “at once” as Locke had it (and therefore making the abstract triangle having incompatible properties), Hume argues that each idea of a particular triangle can elicit in our mind a number of ideas of other triangles. When he says “some ideas are particular in their nature but general in their representation” (T 1.1.7.10), the power of “general ... representation” that an idea may have is nothing but the mental process of associating the one particular idea with other particular ideas with similar contents – i.e., association by resemblance (see Topic 5 below).

For example, when I think about the German shepherd I saw yesterday in the park, by association I start thinking about another dog I saw two days ago which was a retriever, and then the thought of yet another dog comes to mind but this time it is the idea of the poodle I saw last week, and so on. When I am having a chain of ‘dog’ thoughts, each thought is separate and about a particular dog, and the thoughts go through my mind in a quick succession one after another. But I do *not* have an idea of an ‘abstract’ dog which is a German shepherd, a retriever and a poodle all “at once”.

With his scepticism about abstract ideas thus justified, Hume is able to dismiss other ideas that for Descartes and the rationalists seemed to be innate in human beings. For example, we have no clear idea of the infinite because we have no impression of infinite, whether of infinite division, or of infinitely small quantities. Anything of which we have an impression is finite. For Hume, all our ideas are either copies of impressions, or combinations of ideas that are themselves copied from impressions. This is put very clearly in his *Enquiry concerning Human Understanding*, where the idea of God as being all-good and all-wise is taken to be the result of “augmenting without limit those qualities of goodness and wisdom” (E 2.6). The original impressions and subsequent ideas from which the notion of all-good and all-wise has been derived were nothing other than impressions and ideas of everyday goodness and wisdom. Hume takes it for granted – rather too readily, according to critics – that “a blind man can form no notion of colours, a deaf man of sounds (E 2.7).

Summary on Perceptions as Impression and Ideas

Perceptions (i.e., mental contents) divided into:

(a) **Impressions** – original perceptions, **vivid feel**, further divided into:

- Sensations from the bodily senses - e.g., sight, smell, hearing, taste, touch.
- Passions (i.e., sentiments, feelings, emotions) - e.g., pleasure, pain, attraction, repulsion, love, hatred, esteem, disgust, approval, disapproval.

(b) **Ideas** – derived perceptions, usually **less vivid feel**.

- The “**copy principle**”: impressions are copied as **correspondent** ideas and stored in our **memory**.
 - * E.g., impression of pain (feeling it) ➔ [i.e., generates] idea of pain.
 - * E.g., impression of a horse (seeing one) ➔ idea of a horse.
 - * E.g., impression of a horn (seeing one) ➔ idea of a horn.
- In some (rare) occasions, an idea (e.g., the idea of the taste of a lemon) can become very vivid - so vivid that the idea feels exactly like an impression – i.e., like an original experience.
- Simple vs. Complex impressions and ideas.
- Ideas derived from prior impressions and stored in memory can be joined together in our **imagination** to form new ideas.
 - * E.g., idea of a horse + idea of a horn ➔ idea of a unicorn.
 - * E.g., idea of people being burnt + idea of people being cut into pieces + ... etc. ➔ idea of hell.
- All ideas have their origins in, are traceable back to, or analysable in terms of, some prior impressions, whether it is via memory alone or via the combined operation of both memory and imagination.
 - * Idea of a missing shade of blue (Hume’s own possible counter-example? (see T 1.1.1.10)

5. Association of ideas (Readings: Hume T 1.1.4, T 1.3.6, T 1.3.9, A 35²⁸, and E 3)

What Hume calls the “association of Ideas” is a fundamental operating “principle” (i.e., mechanism) of the human mind. The principle operates in three different ways:

- (a) By **resemblance**: The mind (often) associates together ideas with resembling contents. For example, suppose I know Mary and have an idea of her in my memory. Now, I see a picture of Mary and form an idea of the picture. But the content of idea of the picture of Mary resembles the idea of Mary herself so much so that my mind naturally *moves* to the idea of Mary herself by association of ideas via resemblance.
- (b) By **contiguity** in time or place: The mind (often) associates together ideas of things that have been observed to occur in close temporal or spatial proximity. For example, suppose I used to live in a house next to a park. Now I think about (i.e., have the idea of) my old house. My mind then naturally moves to the idea of the park by association of ideas via contiguity in place. For another example, suppose I have often heard two songs being played one after another. Now I think about (i.e., have the idea of) one of the songs, my mind then naturally moves to the idea of the other song by association of ideas via contiguity in time.
- (c) By **causes and effect**: The mind (most often) associates together ideas of things that have been observed to occur in constant conjunction in the past. For example, in many previous instances I observed a billiard ball moving after being hit by another one. Now, I have see a billiard ball moving towards another one and have an expectation (i.e., an idea) that it will hit the other one. By association of ideas of via the constant conjunction of the similar events I have observed in the past, my mind very naturally moves from the expectation of the second billiard ball being hit (the cause) to the further expectation that it will move (the effect) after being hit.

²⁸ The notation “A 35” refers to Hume’s *Abstract of a Book lately Published - entitled A Treatise of Human Nature*, paragraph 37. This notation system for the text is used through out the Hume lecture materials. The online text is also notated in this way.

6. What Can We Know: two kinds of inquiry and reasoning (Readings: Hume E 4, T 1.3.1 – T 1.3.2, T 1.3.7)

According to Hume, all the objects of human inquiry and reasoning can be divided into two kinds (and only two kinds). They are “relations of idea” on the one hand, which are discoverable by **reason** independent of real existence in the universe, and “matters of fact” on the other, which are discoverable by **experience**. Hume recommends that whenever a question of debate arises, we ask whether it is about relations of ideas or about matters of fact, and then we should proceed to use either reason or experience to address the question or debate.

(a) Relations of Ideas

- Here are examples of various kinds of “relations of ideas”:
 - * Hume’s own example: mathematical relations (e.g., Pythagoras theorem in geometry).
 - * Contemporary example: logical relations (e.g., Modus Ponens – the valid argument form or deductive principle: “If P then Q. P is true. Therefore Q is true”).
 - * Contemporary example: conceptual relations between complex ideas and the simpler ideas in terms of which they are defined (e.g., the complex idea “woman” is defined in terms of, and therefore has conceptual relations to, the simpler ideas “adult”, “human” and “female”).
- True relations of ideas are provable by **reason a priori**, or, in Hume’s words, they “are discoverable by the mere operation of thought, without dependence on what is any where existent in the universe” (E 4.1.1). For Hume, reason can only discover true relations of idea by two means:
 - * **Intuition** or rational intuition (by which we discover axioms, i.e., self-evident truths, about relations of ideas).
 - * **Demonstrative reasoning** – i.e., making deductive inferences (by which we discover new relations of ideas that are inferable from axioms).
- Statements expressing relations of ideas are nowadays often called “analytic statements”. In Hume’s view, (analytic) statements that express true relations of ideas are of *certainty* – i.e., *necessarily* true (meaning: the denial or negation of which would be a self-contradiction).

(b) Matters of Fact

- In Hume’s view, “matters of fact” cannot be proven or demonstrated by reason a priori. Rather, matters of fact are discoverable only via **experience a posteriori** by the following means:
 - * **Observation** – i.e., having impressions via the bodily senses, such as seeing, hearing, tasting, smelling, touching, feeling (by which we discover events and objects in the world and also their qualities).
 - * **Inductive reasoning** (Hume also uses the term “experimental reasoning”²⁹) – i.e., making inductive (or non-deductive, or what he also calls “experimental”) inferences from what we already know from past experience to what we do not yet know because it has not been observed. Inductive inferences can be divided into two kinds:
 - + **Inferences from particular to particular**: The mind often infers from an observed *particular* instance to an unobserved *particular* instance - by the mental principle of “association of ideas” via constant conjunction of “cause and effect” as well as via “resemblance” (see Topic 5 above).
 - For example, given my past observations of the constant conjunction of burning and smoke being produced as **cause and effect**, when I am now outside a house and see smoke coming out of its chimney, I infer that something is burning inside even though I do not see inside. In this case, I make a *causal inference* from a particular effect (observed) to a particular cause (unobserved). Later, when I am inside the house and see that coal is burning in the fireplace, I infer that smoke is coming out the chimney even though I do not see the smoke outside. In this case, I infer from a particular cause (observed) to a particular effect (unobserved).
 - For example, given my past experience of several computers of a certain model by a certain manufacturer, which all broken down shortly after purchase, when I now see a new computer of the same model by the same manufacturer, I infer that this computer will also break down shortly after purchase. In this case, I make *analogical inference* from certain qualities of several particular past instances (observed) to an associated quality attributed the new instance (unobserved) - via the

²⁹ E8.17, E9.6. E12.34,

resemblance between the new to the old instances (e.g., they are all of the same model by the same manufacturer).³⁰

- + **Inferences from particular to general:** By association of ideas via “resemblance”, the mind also often infers *from* a limited number of similar particular instances of constant conjunction of certain events or qualities (observed) *to* a *general law* or principle which it believes to universally apply to all (unobserved) similar instances. We nowadays call this type of inference “inductive generalisation”.
 - For example, from my past experience of a limited number of cases where coal burning and smoke being produced had constant conjunction as **cause and effect**, I now infer the universal causal principle that coal burning *always* produces smoke, which I believe to apply to *all* instances of coal burning.
 - For example, from my past experience of a limited number of cases where computer of a certain model by a certain manufacturer all having broken down shortly after purchase, and by association of ideas via **resemblance**, I now infer the universal principle (which is *not* a causal principle in this case) that computers of this model by this manufacturer *always* break down shortly after purchase.
- Statements expressing what Hume calls “matters of fact” (or “what is any where existent in the universe”) are nowadays called “empirical statements”. In Hume’s view, there is no certainty to any matters of fact. Statements that express real matters of fact are *not necessarily* true (meaning: the denial or negation of which would not be a self-contradiction).
- By “necessary”, Hume means “cannot be imagined to be otherwise without self-contradiction”. Hence, a necessary truth is one which would be self-contradictory to deny or imagine to be false. In general, Hume argues that if it is not self-contradictory to imagine a claim to be false, then the claim itself is not necessarily true, even if it is true. Likewise, if it is not self-contradictory to imagine an event not to occur, then there is no necessity to it, even if it has in fact occurred. Hume’s point here is not whether something is in fact true, but whether there is any necessity to its truth if it is indeed true. Indeed, it seems that Hume takes unintelligibility, unimaginability, or inconceivability as an indicator of (if not co-extensive with) impossibility. For example, he writes:

Wherein consists the difference betwixt believing and disbelieving any proposition? The answer is easy with regard to propositions, that are prov’d by intuition or demonstration. In that case, the person, who assents, not only conceives the ideas according to the proposition, but is necessarily determin’d to conceive them in that particular manner, either immediately or by the interposition of other ideas. Whatever is absurd is unintelligible; nor is it possible for the imagination to conceive any thing contrary to a demonstration.³¹
- *
 - * Regarding causal connection between events in particular, Hume argues that it is not a self-contradiction to deny any causal claim in the form “events of type A are always followed by event of type B”. For each event is a distinct existence, Hume argues, and the idea of one event (e.g., the idea “the billiard ball is hit”) is completely separate in its content from the idea of another event (e.g., the idea “the billiard ball moves”) – i.e., there is no conceptual relation between the two ideas. Hence, it is not self-contradictory to imagine a situation where an event of type A occurs but no event of type B follows. So, a causal claim about what events will follow from what events, even if true, is not a necessary truth. Causal claims, if true, are merely *contingent* truths.
- According to Hume, so-called “causal inferences” are nothing more than mere habits of the mind operating in accordance to the principle of “association of ideas” (see Topic 5 above). To say that events of one type “cause” events of another type is nothing more than saying that events of the first type are “constantly followed” by events of the second type. The belief in a causal law (or a “law of nature”) is nothing more than a belief that the future will resemble the past in those regards stated by the law. The mind develops beliefs (i.e., ideas) about cause and effect by the natural operation of the principle of “association of ideas”. It also develops a confidence in those beliefs as the result of the many repeated operations of the principle in the past. But whether the future will always resemble the past, or whether events of one kind will always be follow by events of another kind (which belong to the category of “matters of fact”), can never be “proven” by reason with certainty. Therefore, our confidence in such beliefs about cause and effect is neither warranted by demonstrative reasoning or rational intuition.

³⁰ See Topic 7 below for more details on analogical inference.

³¹ T1.3.7.3.

- “[A]ll probable reasoning is nothing but a species of sensation [...] Objects have no discoverable connexion together; nor is it from any other principle but custom operating upon the imagination, that we can draw any inference from the appearance of one to the existence of another.” (T 1.3.8.12)
- “Those philosophers, who have divided human reason into *knowledge and probability*, and have defin’d the first to be *that evidence, which arises from the comparison of ideas*, are oblig’d to comprehend all our arguments from causes or effects under the general term of probability. But tho’ every one be free to use his terms in what sense he pleases; and accordingly in the precedent part of this discourse, I have follow’d this method of expression; ’tis however certain, that in common discourse we readily affirm, that many arguments from causation exceed probability, and may be receiv’d as a superior kind of evidence. One wou’d appear ridiculous, who wou’d say, that ’tis only probable the sun will rise to-morrow, or that all men must dye; tho’ ’tis plain we have no further assurance of these facts, than what experience affords us. For this reason, ’twould perhaps be more convenient, in order at once to preserve the common signification of words, and mark the several degrees of evidence, to distinguish human reason into three kinds, viz. *that from knowledge, from proofs, and from probabilities*. By knowledge, I mean the assurance arising from the comparison of ideas. By proofs, those arguments, which are deriv’d from the relation of cause and effect, and which are entirely free from doubt and uncertainty. By probability, that evidence, which is still attended with uncertainty. ’Tis this last species of reasoning, I proceed to examine.” (T.1.3.11.2)

7. Sympathy & Moral Sentiments (Readings: Hume T 2.1.11, T 2.2.4-5, T 3.2.2, T 3.3.1-2)³²

“Sympathy” (or what is now often called “empathy”) is in Hume’s view a complex **mechanism** or principle of the human mind which relies on the **combined operation** of three more fundamental principles, including (a) the “copy principle”, (b) the principle of “association of ideas” by resemblance, by contiguity in time or place, and by the constant conjunction of cause and effect, and (c) the principle of more vivid perceptions “enlivening” less vivid associated perceptions.

- (a) The “**copy principle**” (recall Hume Lectures 1 & 2 above): I have had some bad experienced in the past where I felt sad, frustrated, hurt, or angry, and I cried and screamed. These direct experiences or impressions got “copied” as correspondent ideas in my mind and stored in my memory.
- (b) The principle of “**association of ideas**” (again recall Hume Lectures 1 & 2 above):
- After I have had repeated experience of the similar kind, my mind has naturally associated together the various ideas derived from the impressions involved in the experience. For example, by the principle of association of ideas via constant conjunction of **cause** and **effect**, my mind has established a pattern of association between the idea of crying and the idea of sadness. Suppose now in a new instance I observe a person crying in front of me. By the same principle of association of ideas, my mind then naturally moves from the ideas of crying to the **idea of sadness**. In short, due to my own past similar experience, I infer from what I observe to be the outward behaviour of the person (crying) to the existence of an inner state of the person (sadness) which I believe to have caused the person’s behaviour.
 - According to Hume, every human creature resembles ourselves, and by that means has an advantage above any other object in influencing the operation of the principle of association of ideas. So, due to the **resemblance** between myself and the person in front of me (as we are both human beings), my mind, by association of similar ideas, naturally moves from the idea of the person to the **idea of myself**.
 - Furthermore, according to Hume, whenever I think about (i.e., have an idea of) myself, I will also become intimately aware of my various immediate experiences or impressions (e.g., impressions of where my hands are, how my body is feeling at various places). To have an impression of oneself, in Hume’s view, is nothing more than having a chain of various very vivid immediate experiences or impressions. So, from having the idea of myself I now come to have an **impression of myself**, which Hume argues, is always **very vivid**, for it is constituted by the various very vivid immediate experiences or impressions that I have when I reflect on myself.
 - In short, by observing the crying person in front of me, my mind is now entertaining two perceptions: (a) the idea of sadness and (b) the very vivid impression of myself. The two perceptions are now associated together in my mind via **contiguity** in time – as I am experiencing them in very close temporal proximity with each other.
- (c) The principle of **more vivid perceptions “enlivening” less vivid associated perceptions**: According to Hume, it is a fundamental principle of the mind that if a less vivid perception (i.e., an idea) happens to be associated in the mind with a very vivid perception (i.e., an impression), for example, if the two perceptions happen to occur in the mind at about the same time (i.e., contiguity in time), then the very vivid one will “enliven” the less vivid one to such an extent that the latter will become as vivid as the former and therefore turn into a lively impression. After all, the main difference between an impression and a correspondent idea, in Hume’s theory, consists mainly in their different degrees of vividness. A very vividly experienced idea is qualitatively no different from a correspondent impression. In short, by the principle of more vivid perceptions “enlivening” less vivid associated ones, an originally rather faint idea can become so “enlivened” and vivid and forceful that it qualifies as an impression.
- Now, let us go back to the crying person in front of me. In step 2 above, due to the various operations of principle of association of ideas when I am observing the person, my mind has associated together (a) the idea of sadness and (b) the very vivid impression of myself. Since the impression of myself, as we have seen, is so vivid, it “enlivens” the associated idea of sadness so much so that the idea of sadness is “converted” into an **impression (i.e., direct experience) of sadness**. In other words, I now **feel sad** myself. Or in Hume’s word, I

³² For a fuller account on the various principles involved in the operation of sympathy, also see T 1.1.1.1-7, T 1.1.2-3, T 1.1.4.1-2, T 1.3.6.13, T 1.3.8.2, T 1.3.10.3, T 1.3.14.28, T 1.3.15.1-5, T 1.4.6.4, T 2.1.4.2, T 2.1.11.4-8, T 2.2.4.7, T 3.3.1, T 3.3.2.1-7

“partake of” the crying person’s sadness “by sympathy”.

As we have seen, in Hume’s account of sympathy, its operation of in part relies on the association of ideas via “resemblance” (e.g., when the idea of another person excites the idea of myself due to the similarities that I believe to hold between the person and myself). From this, Hume further argues that the greater the **number of similar qualities** (e.g., physical features, characters traits, temperaments, beliefs, values, sentiments, and circumstances) and the greater the **degree of resemblance** that one believes another person to share with oneself, the more strongly one would associate the idea of the person with the idea of oneself, and therefore, other things being equal, the more likely that one would sympathize with the person.

- **Argument by analogy** (or analogical reasoning) is, for Hume, nothing more than the operation of the principle of “association of ideas” via **resemblance**. Suppose we have in many past instances observed objects A1, A2, A3, ... all having the qualities Fs and also the quality K. Suppose now we observe a new object B which has the same qualities Fs. Our mind, by association of ideas via resemblance, will often naturally move to (i.e., infer) the idea that B also has the same quality K. In Hume’s words,

“All our reasonings concerning matter of fact are founded on a species of Analogy. [...] But where the objects have not so exact a similarity, the analogy is less perfect, and the inference is less conclusive; though still it has some force, in proportion to the degree of similarity and resemblance.”³³

- Given Hume’s account of sympathy, one possible explanation for cases where one does not sympathize, or **fails to sympathize**, with another person is that one does not think the other person is similar to oneself. Such “disassociation” or “hyper-separation”³⁴ (let’s call it that) occurs, for example, when a Nazi failed to sympathize with the Jews in a concentration camp before and during WWII, when a slave trader fails to sympathize with the slaves from selling whom he is profiting, when a patriarch fails to sympathize with his young daughter whom he is forcing into marriage, or when a well off person living in a rich and peaceful country fails to sympathize with a refugee seeking asylum (and you can find more examples yourself). It is important, however, to note that the fact that one happens not to sympathize with another person because one does not believe the other person to be similar to oneself, this fact, does not by itself make it true that the other person is not in some important ways similar to oneself. Self interest, fear, contempt, hatred, for example, can often cloud people’s observation and judgement so that they fail to observe similarities that in fact exist between themselves and other people. People may also suppress their sympathy by actively ignoring, by not thinking about, ways in which they are similar to others – for fear that the similarities, once recognized, would excite certain unpleasant sentiments (e.g., distress) that they may not want to experience or share by sympathy.

³³ E 9.1.

³⁴ For a very good account of how “hyper-separation” may occur, see Val Plumwood, *Feminism and the Mastery of Nature*, Routledge 1993.

8. Animals & Nature: from the Humean point of view (Readings: Hume E 9, T 1.3.16, T 2.1.12, T 2.2.12, T 2.3.10.32)³⁵

According to Hume, many nonhuman animals (or beings whom he sometimes calls “sensible creatures”³⁶) **resemble** us in both body and mind. Because of the similarities between us and them, we are able to sympathize with animals in approximate ways that we sympathize with other human beings. That is, we sympathy with animals through the same set of mental principles through which we sympathy with other people (see Topic 7 above).

The Body: Below are some passages where Hume discusses the similarities between human and animal bodies:

“The anatomical observations, formed upon one animal, are, by this species of reasoning [i.e., analogical reasoning, on which causal reasoning also relies], extended to all animals”.³⁷

“’Tis usual with anatomists to join their observations and experiments on human bodies to those on beasts, and from the agreement of these experiments to derive an additional argument for any particular hypothesis. [...] where the structure of parts in brutes is the same as in men, and the operation of these parts also the same, the causes of that operation cannot be different, and that whatever we discover to be true of the one species, may be concluded without hesitation to be certain of the other. [...] as the structure of the veins and muscles, the fabric and situation of the heart, of the lungs, the stomach, the liver and other parts, are the same or nearly the same in all animals, the very same hypothesis, which in one species explains muscular motion, the progress of the chyle, the circulation of the blood, must be applicable to every one; and according as it agrees or disagrees with the experiments we may make in any species of creatures, we may draw a proof of its truth or falshood on the whole.”³⁸

“These *analogical observations* may be carried farther, even to this science [i.e., the science of the human mind], of which we are now treating; and any theory, by which we explain the operations of the *understanding*, or the origin and connexion of the *passions* in man, will acquire additional authority, if we find, that the same theory is requisite to explain the same phaenomena in all other animals.”³⁹

“Let us, therefore, apply this method of enquiry, which is found so just and useful in reasonings concerning the body, to our present anatomy of the *mind*, and see what discoveries we can make by it.”⁴⁰

In the above passages, Hume is in effect putting forward an **argument by analogy** – on how we should reason about the similarities that we observe between humans and animals. His analogical argument can be summarized as follows:

- Premise. To the extent that humans and animals are similar in their **bodily operations**, whatever **anatomical principles/hypotheses** can explain those operations in one case are also applicable to the other case.
- Conclusion. To the extent that humans and animals are similar in their **behaviours**, whatever **mental principles/hypotheses** (e.g., the copy principle, the principle of “association of ideas”, ..., etc.) can explain those behaviours in one case are also applicable to the other case. (analogically inferred)

In short, if we justified in thinking animal bodies and the human body operate in similar ways because of the resemblance in anatomy between animal bodies and the human body, then, Hume argues by analogy, we are also justified to think that animal minds and the human mind operate in some similar ways - given the resemblance in behaviours between humans and animals.

³⁵ For a fuller account on the various principles involved in the operation of sympathy, also see T 1.1.1.1-7, T 1.1.2-3, T 1.1.4.1-2, T 1.3.6.13, T 1.3.8.2, T 1.3.10.3, T 1.3.14.28, T 1.3.15.1-5, T 1.4.6.4, T 2.1.4.2, T 2.1.11.4-8, T 2.2.4.7, T 3.3.1, T 3.3.2.1-7.

³⁶ E 12.1.7.

³⁷ E 9.1.

³⁸ T 2.1.12.2.

³⁹ E 9.1, emphases added.

⁴⁰ T 2.1.12.2, emphases added.

The Mind: Hume holds that the minds of animals of many nonhuman species resemble the human mind in both cognitive and emotional capacities.

Animals and the “Understanding”: In terms of cognitive capacities (or what Hume calls the faculty of “understanding”), Hume argues that it is evident from our observations on animal behaviours that animals, like us, are capable of having **perceptions** and **memories**,⁴¹ and so the “**copy principle**” should likewise operate an animal’s mind in pretty much the same way that it operates the human mind. It is also evident, Hume argues, that animals (or at least animals whose **behaviours** resemble those of a human being) are capable of **inductive reasoning** (or what he sometimes call “experimental reasoning”), which is after all under Hume’s account nothing more than the customary or habitual operation of the mental principle of “association of ideas” – although, he admits, inductive reasoning in animals (like that in a small child or a mentally impaired person) is less sophisticated than that in a normal and educated adult human being. The difference here, Hume argues, is one in **degree**, not in type. It should be noted, however, that Hume does not attribute demonstrative or logical reasoning, or rational intuition, to animals.

Below are some passages where Hume discusses the similarities between animals and human beings in their faculty of understanding, and the common mental principles according to which animal and human perceptions (i.e., impressions and ideas) operate:

“[N]o truth appears to me more evident, than that **beasts are endow’d with *thought* and *reason* as well as men.** [...] When therefore we see other creatures, in millions of instances, perform like actions, and direct them to like ends, all our principles of [inductive] reason and probability carry us with an invincible force to believe the existence of a like cause. [...] ’Tis from the *resemblance* of the *external actions* of animals to those we ourselves perform, that we judge their *internal likewise to resemble* ours; and the same principle of [analogical and causal] reasoning, carry’d one step farther, will make us conclude that since our internal actions resemble each other, the causes, from which they are deriv’d, must also be resembling. When any hypothesis, therefore, is advanc’d to explain a *mental* operation, which is common to men and beasts, we must apply the same hypothesis to both”⁴²

“It seems evident, that animals, as well as men *learn* many things from *experience*, and *infer*, that the same events will always follow from the same causes. By this principle [of association of ideas via constant conjunction of cause and effect] they become acquainted with the more obvious properties of external objects, and gradually, from their birth, treasure up a knowledge of the nature of fire, water, earth, stones, heights, depths, &c. and of the effects, which result from their operation. The ignorance and inexperience of the young are here plainly distinguishable from the cunning and sagacity of the old, who have learned, by long *observation*, to avoid what hurt them, and to pursue what gave ease or pleasure. [...] In all these cases, we may observe, that **the animal infers some fact beyond what immediately strikes his senses;** and that this inference is altogether founded on past experience, while the creature *expects* from the present object the same consequences, which it has always found in its observation to result from similar objects.”⁴³

“A dog, that has hid a bone, often forgets the place; but when brought to it, his thought passes easily to what he formerly conceal’d, by means of the *contiguity*, which produces a relation among his ideas. In like manner, when he has been heartily beat in any place, he will tremble on his approach to it, even tho’ he discover no signs of any present danger. The effects of *resemblance* are not so remarkable; but as that relation makes a considerable ingredient in *causation*, of which all animals shew so evident a judgement, we may conclude that the three relations of resemblance, contiguity and causation operate in the same manner upon beasts as upon human creatures.”⁴⁴

“Animals [...] are not guided in these inferences by [demonstrative] reasoning: Neither are children: Neither are the generality of mankind, in their ordinary actions and conclusions: Neither are philosophers themselves, who, in all the active parts of life, are, in the main, the same with the vulgar, and are governed by the same maxims. [...] the *experimental reasoning* itself, which we possess in common with beasts, and on which the whole conduct of life depends, is nothing but a *species of instinct or mechanical power*, that acts in us unknown to ourselves; and in its chief operations, is not directed by any such relations or comparisons of ideas, as are the proper objects of our intellectual faculties. Though the instinct be different, yet still it is an instinct, which teaches a man to avoid the fire; as much as that, which teaches a bird, with such exactness, the art of incubation, and the whole economy and order of its nursery.”⁴⁵

⁴¹ T 1.3.16.6-7.

⁴² T 1.3.16.1-3, emphases added.

⁴³ E 9.2-4, emphases added.

⁴⁴ T 2.1.12.7, emphases added.

⁴⁵ E 9.5-6, emphases added.

Animals and the “Passions”: According to Hume, many animals, like us, are capable of having a great variety of feelings, sentiments, and passions, such as pleasure and pain, love and hatred, pride and humility, fear and grief. Moreover, Hume thinks that animals are capable of **sympathy** too!

Below are some passages where Hume discusses the similarities between animals and human beings in their passions and the common mental/psychological principles according to which their passions operate:

“We must first shew the correspondence of *passions* in men and animals, and afterwards compare the *causes*, which produce these passions. [...] every species of creatures, which approach so often to man, as to familiarize themselves with him, show an evident pride [a passion] in his approbation [the cause of the passion], and are pleas’d [a passion] with his praises and caresses [the causes] [...] pride and humility are not merely human passions, but extend themselves over the whole animal creation. The *causes* of these passions are likewise much the same in beasts as in us, making a just allowance for our superior knowledge and understanding.”⁴⁶

“The next question is, whether, since those passions are the same, and arise from the same causes thro’ the whole creation, the manner, in which [i.e., the mental principles, by which] the causes operate, be also the same. According to all rules of analogy, this is justly to be expected”.⁴⁷

“There are [...] instances of the relation [or association] of impressions, sufficient to convince us, that there is an union of certain affections with each other in the inferior species of creatures as well as in the superior, and that their minds are frequently convey’d thro’ a series of connected emotions. A dog, when elevated with joy, runs naturally into love and kindness, whether of his master or of the sex. In like manner, when full of pain and sorrow, he becomes quarrelsome and ill-natur’d; and that passion, which at first was grief, is by the smallest occasion converted into anger [by the principle of union among similar passions]. [...] all the internal principles, that are necessary in us to produce [the passions], are common to all creatures; and since the causes, which excite these passions, are likewise the same, we may justly conclude, that these causes operate after the same manner [i.e., by the same mental principles] thro’ the whole animal creation.”⁴⁸

“But to pass from the passions of love and hatred [...] as they appear in man, to the same affections, as they display themselves in brutes; we may observe, not only that love and hatred are common to the whole sensitive creation, but likewise that their causes [...] are of so simple a nature, that they may easily be suppos’d to operate on mere animals. [...] As animals are but little susceptible either of the pleasures or pains of the imagination, they can judge of objects only by the sensible good or evil, which they produce, and from that must regulate their affections towards them. Accordingly we find, that by benefits or injuries we produce their love or hatred; and that by feeding and cherishing any animal, we quickly acquire his affections; as by beating and abusing him we never fail to draw on us his enmity and ill-will.”⁴⁹

“Love in beasts is not caus’d so much by relation, as in our species; and that because their thoughts are not so active as to trace relations, except in very obvious instances. Yet ’tis easy to remark, that on some occasions it has a considerable influence upon them. Thus acquaintance, which has the same effect as relation, always produces love in animals either to men or to each other. For the same reason any likeness among them is the source of affection. An ox confin’d to a park with horses, will naturally join their company, if I may so speak, but always leaves it to enjoy that of his own species, where he has the choice of both.”⁵⁰

“’Tis evident, that sympathy, or the communication of passions, takes place among animals, no less than among men. Fear, anger, courage and other affections are frequently communicated from one animal to another, without their knowledge of that cause, which produc’d the original passion. Grief likewise is receiv’d by sympathy; and produces almost all the same consequences, and excites the same emotions as in our species. The howlings and lamentations of a dog produce a sensible concern in his fellows. And ’tis remarkable, that tho’ almost all animals use in play the same member, and nearly the same action as in fighting; a lion, a tyger, a cat their paws; an ox his horns; a dog his teeth; a horse his heels: Yet they most carefully avoid harming their companion, even tho’ they have nothing to fear from his resentment; which is an evident proof of the sense brutes have of each other’s pain and pleasure.”⁵¹

⁴⁶ T 2.1.12.3-5, emphases original.

⁴⁷ T 2.1.12.6-9, first emphasis original, the rest added.

⁴⁸ T 2.1.12.8-9, first emphasis added, the rest original.

⁴⁹ T 2.2.12.1, T 2.2.12.3, emphasis original.

⁵⁰ T 2.2.12.4.

⁵¹ T. 2.2.12.6.

In short, Hume argues that from observations of animal **behaviours**, it is evident that animals (or mammals anyway) resemble human beings to a significant extent in their faculty of **understanding** (or cognition) – including their capacities for (i) *perceptions*, (ii) *memories*, and for (iii) *inductive reasoning* by association of ideas via resemblance, contiguity, and the constant conjunction of cause and effect. Hume also argues that many animals resemble humans in their faculty of the **passions** (or what he calls the faculty of “taste”) – including their capacities for (iv) experiencing a great variety of *feelings and emotions*, the operations of which in animals are guided by similar principles as in humans, and their capacity for (v) *sympathy* – not only with members of their own species, but also with members of other animal species.

The question now is: If even animals, with inferior understanding, are capable of sympathizing with members of another animal species, then shouldn't we, human beings, with superior understanding, also be capable of sympathizing with animals beyond the boundary of our own species?

Sympathy, on Hume's account, operates partly via the principle of association of idea by **resemblance**. We sympathize with those who we believe to be similar to us. By sympathy, their emotions and passions communicate to us in so lively a manner that we come to share and experience those emotions and passions ourselves. Given Hume's account of similarities between humans and animals, his account of sympathy, and his account of the human understanding (which is much more sophisticated than that in animals and therefore is much better in discovering the resemblance between humans and animals), the answer to the question is clear: We are indeed equipped with the capacities necessary and sufficient for developing and having sympathy with nonhuman animals.

When we fail to sympathize with an animal in pain or in distress, for example, this failure in sympathy, under Hume's account, is rather similar to the failure to sympathize with a fellow human being. Both cases can be explained in terms of factors that may block the operations of the various mental principles on which sympathy relies – blocking factors such as self-interest, fear, or feeling of superiority, or general insensitivity, or even autism spectrum disorders.

But in both the human and nonhuman case, failure of sympathy can be mended to some extent by the better use of the human understanding, which should enable us to see not only the differences, but also the similarities, between ourselves and other people as well as other sentient animals.

However, if people were not naturally borne with the capacities to detect the emotions of others, and to reproduce the same kinds of emotions in themselves, then no amount of reasoning and the exercise of their cognitive powers will enable them to sympathize with others. At its core, sympathy is an emotional capacity to **reproduce passions** of others in oneself. It plays an essential role in our moral judgements. Without the ability to sympathize, our moral judgements would be as unreliable as a colour-blind person's aesthetic judgement of a painting.

Reason can present us with facts about the world, make clear distinctions that help us avoid confusions and inconsistency. But after all of that is done, it is for the passions produced by sympathy to take the role of the final judge of moral good and evil, right and wrong. I take this to be what Hume means in part – when he says: **“Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them”**.⁵²

⁵² T 2.3.3.4.