

Photocopiable/digital resources may only be copied by the purchasing institution on a single site and for their own use

2017 specification

first exams in 2019 (2018 for AS)

Course Companion for   
AS / A Level Year 1 AQA

Paper 1, Section B: Moral Philosophy

Philosophy AS / A Level | AQA | 7171/7172

zigzageducation.co.uk

Publish your own work... Write to a brief...

Register atpublishmenow.co.uk

POD

9765

🖰 Follow us on Twitter **@ZigZagRS**

**Contents**

[Thank You for Choosing ZigZag Education ii](#_Toc11077563)

[Teacher Feedback Opportunity iii](#_Toc11077564)

[Terms and Conditions of Use iv](#_Toc11077565)

[Teacher’s Introduction 1](#_Toc11077566)

[Key Terminology in Moral Philosophy 2](#_Toc11077567)

[Normative Ethical Theories 3](#_Toc11077568)

[Introduction to Normative Ethical Theories 3](#_Toc11077569)

[Ethical Theories 1: Utilitarianism 5](#_Toc11077570)

[Bentham’s Utilitarianism 5](#_Toc11077571)

[Mill’s Utilitarianism 8](#_Toc11077572)

[*Higher/Lower Pleasures* 12](#_Toc11077573)

[Act and Rule Utilitarianism 14](#_Toc11077574)

[The Strengths of Hedonistic Utilitarianism 16](#_Toc11077575)

[Issues for Hedonistic Utilitarianism 17](#_Toc11077576)

[Preference Utilitarianism 26](#_Toc11077577)

[Issues with Preference Utilitarianism 28](#_Toc11077578)

[Ethical Theories 2: Kantian Deontological Ethics 30](#_Toc11077579)

[Introduction to Kant 30](#_Toc11077580)

[The Good Will 31](#_Toc11077581)

[Duty, Reason and Self-interest 33](#_Toc11077582)

[1: Categorical and Hypothetical Imperatives 34](#_Toc11077583)

[2: First Formulation of the Categorical Imperative 35](#_Toc11077584)

[3: Second Formulation of the Categorical Imperative 37](#_Toc11077585)

[Strengths of Kantian Ethics 38](#_Toc11077586)

[Issues with Kantian Deontological Ethics 39](#_Toc11077587)

[Ethical Theories 3: Aristotle’s Virtue Ethics 45](#_Toc11077588)

[Introduction to Virtue Ethics 45](#_Toc11077589)

[Aristotle and the *Nicomachean Ethics* 46](#_Toc11077590)

[1: What is ‘the Good’? 46](#_Toc11077591)

[2: The Function Argument 48](#_Toc11077592)

[1. The good for human beings is performing their functions well. 49](#_Toc11077593)

[2. The characteristic human function is ‘an activity of the soul which follows or implies a rational principle’. 50](#_Toc11077594)

[3: Virtues and the Doctrine of the Mean 51](#_Toc11077595)

[4: Moral Responsibility 56](#_Toc11077596)

[5: Practical Reasoning, Contemplation and Pleasure 59](#_Toc11077597)

[The Strengths of Virtue Ethics 61](#_Toc11077598)

[Issues with Aristotle’s Virtue Ethics 63](#_Toc11077599)

[Applied Ethics 68](#_Toc11077600)

[Introduction to Applied Ethics – Use of Utilitarianism, Kantian Ethics and Virtue Ethics 68](#_Toc11077601)

[Applied Ethics 1: Stealing 68](#_Toc11077602)

[Applied Ethics 2: Simulated Killing 71](#_Toc11077603)

[Applied Ethics 3: Eating Animals 75](#_Toc11077604)

[The Utilitarian Perspective 75](#_Toc11077605)

[The Kantian Perspective 76](#_Toc11077606)

[The Aristotelian Perspective 77](#_Toc11077607)

[Applied Ethics 4: Lying 78](#_Toc11077608)

[Meta-ethics 81](#_Toc11077609)

[Introduction to Meta-ethics 81](#_Toc11077610)

[Ethical Language 1: Cognitivism 82](#_Toc11077611)

[Moral Realism 84](#_Toc11077612)

[Ethical Naturalism 84](#_Toc11077613)

[Moore’s Intuitionism 87](#_Toc11077614)

[Moral Anti-realism 91](#_Toc11077615)

[Error Theory 91](#_Toc11077616)

[Ethical Language 2: Non-cognitivism 93](#_Toc11077617)

[Emotivism 93](#_Toc11077618)

[Ayer’s Emotivism 94](#_Toc11077619)

[Prescriptivism 96](#_Toc11077620)

[Issues for Moral Anti-realism 98](#_Toc11077621)

[Appendices 102](#_Toc11077622)

[Appendix 1: Online Resources for Teachers and Students 102](#_Toc11077623)

[Appendix 2: Reader’s Guide to the Anthology Texts 103](#_Toc11077624)

[Appendix 3: A Brief Guide to Writing Philosophy 105](#_Toc11077625)

[Answers 106](#_Toc11077626)

[Answers to Activities 106](#_Toc11077627)

[Answers to Quick Quizzes 108](#_Toc11077628)

Thank You for Choosing ZigZag Education

Thank you

for choosing ZigZag Education!

Love it as it is?

Let the author and other teachers know what you think

Got a suggestion?

If your improvement leads to an update we will send you a new copy for free

Found a problem?

We will fix it and send you a free updated copy

We ♥ your feedback!

Let us know what you think using the feedback sheet on the next page.

£10 ZigZag Voucher for detailed & complete reviews!

Talk to Us!

Post:

ZigZag Education, Unit 3, Greenway Business Centre, Doncaster Road, Bristol BS10 5PY



Fax:

0117 959 1695



Real Person:

0117 950 3199



Web:

zzed.uk/more



Email:

philosophy

@zigzageducation.co.uk



For more resources go to zzed.uk/more

where you can preview before you buy

ZigZag is a large community of over 6000 teachers & educationalists.

Review new titles, publish your own work or write to a brief.

Fancy being involved?

Then register at…

publishmenow.co.uk

*The Professional Publishing Community*

Alternatively email new resource ideas directly to…

publishmenow@zigzageducation.co.uk

Become a Published Author

Teacher Feedback Opportunity

**Teacher Feedback Opportunity**

£10 ZigZag Voucher for detailed & complete reviews! ⬩ Use for problems/areas for improvement/positive feedback

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Resource ID & Name** | | | 9765 Course Companion for AS / A Level Year 1 AQA Philosophy: Moral Philosophy | | | |
|  | | | | | | |
| **School Name** | |  | | | | |
|  | | | | | | |
| **Your Name** |  | | |  | **Position** |  |

Overall, what did you think about this resource?

I particularly like this resource because…

How does it help you or your students?

It is better than some other resources because…

What might you say to a colleague in a neighbouring school to persuade them to use this resource?

How well does it match your specification (and which specification is this)?

Other comments, suggestions for improvements, errors found (please give page numbers) etc.

🞏✓ Please email details of new Philosophy resources. Email address:

|  |  |
| --- | --- |
| *Resources* I would like published: |  |
|  |  |
| *Resources* I might write*, or have written,* for consideration for publication: |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Fax to: | Email to: | Submit online: | Post to: |  | INTERNAL USE ONLY  Feedback logged: ✓🞎  Complete & detailed: Y /N  If detailed, £10 sent: ✓🞎 |
| **0117 959 1695** | **feedback@**  **zigzageducation.co.uk** | **zzed.uk/feedback** | **ZigZag Education, Unit 3, Greenway Business Centre, Doncaster Road, Bristol BS10 5PY** |  |

**Terms and Conditions of Use**

**Terms and Conditions**

Please note that the **Terms and Conditions** of this resource include point 5.3, which states:

|  |
| --- |
| **‘You acknowledge that you rely on your own skill and judgement in determining the suitability of the Goods for any particular purpose.’** |

‘We do not warrant: that any of the Goods are suitable for any particular purpose (e.g. any particular qualification), or the results that may be obtained from the use of any publication, or expected exam grades, or that we are affiliated with any educational institution, or that any publication is authorised by, associated with, sponsored by or endorsed by any educational institution.’

**Copyright Information**

Every effort is made to ensure that the information provided in this publication is accurate and up to date but no legal responsibility is accepted for any errors, omissions or misleading statements. It is ZigZag Education’s policy to obtain permission for any copyright material in their publications. The publishers will be glad to make suitable arrangements with any copyright holders whom it has not been possible to contact.

Students and teachers may not use any material or content contained herein and incorporate it into a body of work without referencing/acknowledging the source of the material (‘Plagiarism’).

**Disclaimers**

This publication is designed to supplement teaching only. Practice questions may be designed to follow the content of a specification and may also attempt to prepare students for the type of questions they will meet in the examination, but will not attempt to predict future examination questions. ZigZag Education do not make any warranty as to the results that may be obtained from the use of this publication, or as to the accuracy, reliability or content of the publication.

Where the teacher uses any of the material from this resource to support examinations or similar then the teacher must ensure that they are happy with the level of information and support provided pertaining to their personal point of view and to the constraints of the specification and to others involved in the delivery of the course. It is considered essential that the teacher adapt, extend and/or censor any parts of the contained material to suit their needs, the needs of the specification and the needs of the individual or group concerned. As such, the teacher must determine which parts of the material, if any, to provide to the students and which parts to use as background information for themselves. Likewise, the teacher must determine what additional material is required to cover all points on the specification and to cover each specification point to the correct depth.

ZigZag Education is not affiliated with Pearson, Edexcel, OCR, AQA, WJEC, Eduqas, SQA, CCEA, CIE, International Baccalaureate Organization or DFE in any way nor is this publication authorised by, associated with, sponsored by or endorsed by these institutions unless explicitly stated on the front cover of this publication.

Links to other websites, and contextual links are provided where appropriate in ZigZag Education publications. ZigZag Education is not responsible for information on sites that it does not manage, nor can we guarantee, represent or warrant that the content contained in the sites is accurate, legal and inoffensive, nor should a website address or the inclusion of a hyperlink be taken to mean endorsement by ZigZag Education of the site to which it points. This includes websites that users are directed to via the convenient zzed.uk short URLs.

**Acknowledgements**

The following images are licensed under Creative Commons Attribution 2.0 License. These are reused and distributed under the terms and conditions found at: [**http://creativecommons.org/licenses/by/2.0/**](http://creativecommons.org/licenses/by/2.0/)

* Image of the 1920 Icall Machine

**Teacher’s Introduction**

Register your email address to receive any future free updates\* made to this resource or other Philosophy resources your school has purchased, and details of any promotions for your subject.

*\* resulting from minor specification changes, suggestions from teachers   
and peer reviews, or occasional errors reported by customers*

Free Updates!

Go to zzed.uk/freeupdates

The AQA AS and A Level Philosophy courses are at times extremely challenging, making the subject both difficult to teach and difficult to learn. My chief aim in writing this course companion for Paper 1, Section B: Moral Philosophy has been to create something which is more student-friendly than the available textbooks, while not sacrificing the nuance and precision required to achieve the top grades. To this end, both detail and accessibility are emphasised, with simple concepts and proofs building into more detailed critiques of philosophers’ works. The most able and passionate students will likely read the anthology texts for themselves.

Each section begins with either a brief introductory paragraph or a thought experiment that builds into more complex ideas and ethical dilemmas. The bulk of this resource consists of expositions of the main arguments followed by their critiques. Sometimes these are presented premise by premise, sometimes I have felt it better to express them in continuous prose. Interspersed throughout are opportunities for discussion, group tasks or written work. The aim of these activities is threefold: (i) to deepen understanding, (ii) to prompt critical thinking and (iii) to reinforce learning.

At the end of the resource you will find several appendices, including a glossary of key terms, a guide to online resources and a guide to anthology texts which should assist and clarify further study.

I wish both yourself and your students the very best of luck in the weeks and months ahead.

*A web page containing all the links listed in this resource is conveniently provided on ZigZag Education’s website at* **zzed.uk/*9765***

🖰

*You may find this helpful for accessing the websites rather than typing in each URL.*

*D Anthony, June 2019*

## Key Terminology in Moral Philosophy

Abc

Like any area of philosophy, ethics has its own set of specialised terminology. The words below will reappear   
time and time again during this course and it is a good idea to know what each means in time for the exam. There may be a lot here to take in right now, some of which you may not presently understand, but if you refer back to this page as you study, you will soon become fluent in the language of ethics.

|  |  |
| --- | --- |
| **Absolute** | Term applied to ethical theories which hold that what is right or wrong will be the same for all people, at all times (e.g. Kantian deontology) |
| **Act-centred** | Term applied to ethical theories which make judgements about actions rather than persons (e.g. utilitarianism and Kantian deontology) |
| **Agent-centred** | Term applied to ethical theories which make judgements about persons rather than actions (e.g. Aristotle’s virtue ethics) |
| **Applied ethics** | The application of ethical thinking to real-world issues, e.g. war or animal rights (the second major topic in this course) |
| **Cognitive** | Language which makes claims about reality that are true or false, i.e. language which states facts |
| **Consequentialist** | Term applied to ethical theories which judge whether an action is right or wrong on the basis of consequence of the actions, not the action itself (e.g. utilitarianism) |
| **Deontological** | Term applied to ethical theories which make judgements about ethical actions based on the intention of the moral agents or whether they are consistent with moral laws (e.g. Kantian deontology) |
| **Descriptive ethics** | A *description* of an individual’s or a group of individuals’ ethical behaviour; in contrast to normative ethics, descriptive ethics is considered, in theory, to be a matter of fact rather than debate, e.g. ‘Most people are selfish’, ‘Most people do not enjoy causing others pain’. Descriptive ethics tell us how things *are,* rather than how they *ought* to be. |
| **Ethics** | The area of philosophy concerned with what behaviour is morally right or wrong, good or bad. |
| **Immoral** | Actions, behaviour or intentions that are considered bad or wrong. |
| **Maxim** | A moral rule. |
| **Meta-ethics** | The branch of ethics which discusses what ethical language means. Typical questions include, ‘What does it mean to say something is good?’, ‘Does ethical language refer to anything in objective reality?’ (the third major topic in this course) |
| **Moral** | Actions, behaviour or intentions that are considered good or right |
| **Moral agent** | Someone involved in making ethical or moral decisions |
| **Moral/ethical theory** | A set of ideas about ethics and how people should behave |
| **Moral realism** | The view that mind-independent moral properties and facts exist |
| **Moral anti-realism** | The view that mind-independent moral properties and facts do not exist |
| **Non-cognitive** | Language which does *not* make claims about reality that are true or false, i.e. language which does *not* state facts |
| **Normative ethics** | The branch of ethics which discusses what individuals *ought* and *ought not* do. Typical questions include, ‘What moral rules should people follow?’, ‘What is it to be a good person?’ (the first major topic in this course) |
| **Objective** | Something which is part of mind-independent reality; a fact. It is true for all people regardless of age, culture, gender, etc. *Note*: this is not the same as *absolute* – Bentham’s principle of utility is *relativistic* but it is *objective* because it holds that pleasure is a part of mind-independent reality. |
| **Relative** | Term applied to ethical theories which hold that actions which may be right in one circumstance, may be wrong in another, and vice versa (e.g. utilitarianism) |
| **Subjective** | Something which is mind-dependent, e.g. an opinion. May be true for some individuals, but not for others. *Note:* this is not the same as *relative* – subjectivist theories of ethics hold that there is *no such thing* as ethical truths, relativist theories hold that what is true in one situation may be false in another. |

Normative Ethical Theories

## Introduction to Normative Ethical Theories

#### The Trolley Problem

Imagine one day you are walking home from school and for a change decide to walk a route home by the railway tracks. However, as you follow the tracks, you begin to hear a shouting in the distance. Following these shouts you arrive at a ledge overlooking a portion of the track where five people are tied down, and unable to get up. Moreover, a train is rapidly approaching them, and as they spot you they begin desperately calling for help, pointing towards a lever that when pulled will divert the train from the five people tied down on the track. However, diverting the train in this way will direct it towards another portion of the tracks where another person, alone, is tied down and also shouting for help. You shout back to the five people tied down, telling them that pulling the lever will result in the death of the other person, but they justify their argument by pointing out ‘There are five of us and only one of him! By pulling the lever, you will save four more lives than if you refrained from pulling it. Surely that is a better outcome?’

By this time the train is fast approaching and you have to make a decision. What is the correct action in this circumstance?

#### What would you do?

This thought experiment, known as the *Trolley Problem*, was devised by the moral philosopher Philippa Foot in the late 1960s (the trolley referred to in the title is not the kind found outside Tesco but, as in the story, a variety of train). It has been widely used by both philosophers and psychologists as a litmus test of an individual’s ethical instincts. The dilemma it poses gets right to the heart of what ethics is about: when faced with a situation where there is no easy option, what exactly should we do? In the first part of this course, we will be looking at three ways philosophers have tried to answer that question.

Our first port of call is utilitarianism, which holds that, fundamentally, it is *consequences* which really count. As a result it is often referred to as a *consequentialist* theory. If you read the *Trolley Problem* and intuitively felt that you should switch the tracks in order to save five people at the expense of one, then chances are that you have utilitarian leanings. In fact, the majority of people do, but the theory itself found its most sustained treatment in the works of the nineteenth-century English philosophers Jeremy Bentham and John Stuart Mill.

Yet many counterarguments have been offered that would seek to demonstrate that the *Trolley Problem*, as a hypothetical scenario, does not encapsulate completely the way human beings engage morally with the world. Similarly, others have offered suggestions for how to address the *Trolley Problem* without purely referring to the consequences. For, as will be examined, analysing actions based purely on whether they have a good or bad outcome is not always coherent, and can lead to some unappealing and counterintuitive conclusions.

#### Analysing ethical arguments

The *Trolley Problem* in particular forces you to think about how you make moral decisions, and highlights perhaps how a good system of ethics should function. For example, it cannot be the case that when thinking about what to do, we can go through a complex process of evaluation before arriving at a correct moral decision, for many ethical dilemmas such as the *Trolley Problem* are greatly time-sensitive, and, unless moral decisions can be made quickly and effectively, it becomes hard to understand how it might effectively be applied to ordinary life.

Such considerations form part of the difficulty of analysing **normative ethics**; the study of what one ought and ought not to do. There are no clear independent standards of ethics that it is possible to refer to and much of the discussion of ethics relies on examining how human beings ethically engage with the world, and what is our common, intuitive understanding of ethics as a whole. This unclear basis has inspired great philosophical debate, and, as will be focused on later in the syllabus in the Meta-ethics section, this lack of clarity has led some philosophers to claim that ethics itself is meaningless or incoherent, and there are no moral facts about what individuals really ought to do. However, for the moment, when working through the different normative ethical theories, keep the following questions in mind:

 Where does the ethical theory believe human beings should look to derive conclusions about what one ought to do? Is it a particular aspect of human nature or thought, or even some feature of the external world?

 What principles does the ethical theory put forward as the measure by which human beings should judge their actions? Is it their consequences, or whether they follow some specific set of rules?

 How does the ethical theory arrive at its key moral principles? Is its reasoning justified?

 What assumptions does the ethical theory make about human beings and the world? Are these assumptions acceptable?

 Why does the ethical theory believe human beings are motivated to act in accordance with the principles they describe?

 Does the ethical theory cover all the ways human beings would intuitively morally act? Are there reasons to believe that certain common moral intuitions may be wrong?

Ethical Theories 1: Utilitarianism

Bentham’s Utilitarianism

#### What is Utility?

Jeremy Bentham is often considered the founder of the utilitarian philosophy. His groundbreaking *Introduction to the Principles of Morals and Legislation* opens with these famous words:

Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do, as well as to determine what we shall do.

Crucially, Bentham is arguing that being ruled by pain and pleasure is just the way humans are. We have no choice in the matter; we must seek pleasure, and we must avoid pain. He goes on to argue that basing our ethical decision-making (that is, how we decide what to do) on anything else would simply be foolish.

#### The Principle of Utility

**Anthology Text:**

Bentham, J (1789), ‘The Principle of Utility’ in *Introduction to the Principles of Morals and Legislation*



Having made this claim about human nature (these are sometimes called *descriptive claims* – they describe how things are), Bentham goes on to make his *normative claim* (he wants to tell us how things ought to be). This is the *principle of utility* which states:

When faced with an ethical decision, we should choose the course of action which *maximises* pleasure and *minimises* pain for the *greatest number of people*.

Go back to the story. Can you see how choosing to save lots of people maximises pleasure and minimises pain more than the alternative?

#### Hedonism and Utility

The principle of utility above relates what is good and right ultimately to what is pleasurable, a theory called **hedonism**. However, it is important to separate the different hedonist claims that Bentham makes. While Bentham makes the descriptive claim that human beings are naturally motivated to act according to what is pleasurable and against what is painful, he makes a different claim when he states that this is what human beings *ought* to be motivated by. As such, the distinctions below can be made between different forms of hedonistic thought.

**Psychological Hedonism –** This is the descriptive, or psychological theory that when human beings act, they are solely motivated by the pursuit of pleasure and the avoidance of pain. This is often used as a foundation for utilitarian thought, yet it has been disputed whether human action and motivation can be reduced down to such a simple claim.

**(Moral) Hedonism** – This is the theory that the right action for an individual is one that generates the most pleasure for themselves, and concurrently the least pain. This expresses, therefore, the belief that one not just simply naturally seeks pleasure, but *ought* to seek pleasure in any situation.

**Hedonistic Utilitarianism** – This is the belief that arises out of Bentham’s principle of utility; that what is right is what generates the most amount of pleasure, not just for oneself, but for everyone a particular action affects.

#### http://images.clipart.com/thw/thw11/CL/5344_2005010018/000803_1076_04/21614792.thb.jpg?000803_1076_0475_v__vDo Utilitarians Have to Accept Psychological Hedonism?

It might seem as if there is a natural fit between utilitarian theories and a belief in psychological hedonism. For it is arguably easier to justify the belief that human beings ought to seek pleasure if all human beings naturally seek pleasure in the world. However, many modern utilitarians do not accept hedonism as the basis of their ethical theories, and have turned towards arguing other concepts, such as happiness, welfare or preferences, should be maximised. Why is this the case?

The first difficulty one could note is that there is a difference between human beings seeking pleasure for themselves, and seeking the greatest amount of pleasure for everybody involved. For example, if I had a chocolate bar, and the choice of eating it myself (generating a moderate amount of pleasure) or giving it to my friend who was fanatical about chocolate (generating a great amount of pleasure), the utilitarian would argue I should give the chocolate bar away. Yet such a decision would conflict with my own desires if psychological hedonism were true, and conflict perhaps even further if I believed in personal hedonism only. Therefore, there arguably is a gap between accepting psychological hedonism, and accepting hedonistic utilitarianism. The former stipulates I perhaps am only motivated by my own pleasure, while the latter believes there are grounds for believing human beings could be motivated to right action by consideration of other people’s pleasure.

The second difficulty potentially arises as a result of the first; namely that am I really only motivated by the pursuit of pleasure? Are there not many altruistic, or sacrificial acts human beings perform that at first glance seem to generate no pleasure for the person performing them? At Halloween, I might gain great pleasure if I ate all the chocolate reserved for trick and treaters, but I still restrain myself and save it for children who knock at the door. A strict psychological hedonist might argue that in restraining myself, I am secretly engaging in an act that gives me a particular form of self-satisfied pleasure, but others would dispute such a conclusion, and argue there are greater rational considerations that take precedence over my own desire to engage in pleasure-generating activities.

Therefore, it is important to initially note that psychological hedonism may not be the best foundation for utilitarianism. Bentham and John Stuart Mill both supported hedonistic versions of utilitarianism, and perhaps don’t completely endorse the **strong** claim that human beings are **only** motivated by pleasure and pain. Rather many have perceived them to simply endorse a **weak** claim; that since human beings are **primarily** motivated by pain and pleasure, such concepts are the most appropriate foundations for any ethical theory. Such claims will be analysed later in Mill’s proof of utilitarianism.

#### Discussion:

*Do you think that human beings are motivated only by pleasure and pain? Can you give examples of when human beings might be motivated by other considerations, or can these all be reduced to matters of pleasure and pain?*

#### What Motivated Bentham to Outline Utilitarianism?

Bentham was greatly concerned with legal issues during his life, and looked partially to develop a moral system that not only was accurate and fair at heart, but could be widely applied to different areas of criminal justice. During the late eighteenth century and early nineteenth century when he lived, there was a significant lack of oversight into crime and punishment. London, where he spent most of his life, had become overpopulated in many places, filled with crime and poverty. Furthermore, due to a lack of a centralised police force and judicial system, punishments were often arbitrary, hard and, in many cases, unjust.

Bentham believed that any legal reform to this system had to have a proper moral system also underpinning it, and from his utilitarian beliefs argued for much social reform, including abolishing slavery, the death penalty and much of corporal punishment which was extensively used at the time. He also developed a specific form of prison, the Panopticon, which, due to a special circular structure, guaranteed all inmates could be potentially observed by a single watchman at any time. However, Bentham’s dedication to these very mechanical solutions to social problems drew much criticism during his life, and many philosophers argue that his version of utilitarianism is similarly reductive, and fails to capture the nuances of moral and social activity. This can be particularly observed in the way Bentham attempts to quantify and analyse pleasure and pain in his utility calculus.

#### How is Utility Calculated? Bentham’s Utility Calculus

Since Bentham was trained as a lawyer and had a keen interest in social and legal reform, it is no surprise that he intended his system to have practical uses. To that end, he devised what has come to be known as the *utility calculus*, a method for determining *quantitatively* (in terms of quantities, i.e. numbers) the right course of action (a hedonist is a person who seeks pleasure above all else; like many other terms, it comes from the Greek, in this case *hedone*, meaning ‘pleasure’).

Bentham lists seven factors which must be taken into account when calculating the actual amount of pleasure an act will produce.

**Act Utilitarianism**

The theory which holds that the right *action* is the one which maximises pleasure   
(or happiness) and   
minimises pain.



1. Its *intensity* (how strong it is)

2. Its *duration* (how long it lasts)

3. Its *certainty* or *uncertainty* (how likely or unlikely it is to actually occur)

4. Its *propinquity* or *remoteness* (how close or far away the pleasure is)

5. Its *fecundity* (how likely it is to lead to more pleasure)

6. Its *purity* (how likely a pleasure or pain is to lead to its opposite; e.g. alcohol may be pleasurable, but it is an impure pleasure as it can lead to its opposite: pain, in the form of a hangover or worse)

7. Its *extent*; that is, the number of persons to whom it *extends*; or (in other words) who are affected by it.

The calculus is a distinctive feature of Bentham’s version of utilitarianism. Significantly, it maintains that ethical decisions should be made on a case-by-case basis; thus it is known as **act utilitarianism**.

Later in the guide, in-depth issues will be explored concerning whether it is possible to calculate pleasure and pain in this manner, but for the moment consider the following questions:

 Is it possible to compare pleasures, let alone quantify them? How could a travel agent judge effectively, for example, whether my friend or I enjoy skiing more when trying to give away a free skiing holiday?

 Does the utility calculus present too many factors for human beings to efficiently evaluate in a time-sensitive ethical dilemma? What if I’m faced with a snap moral decision, could I really assess the fecundity of all my possible actions?

 Is it important to only assess the quantity of pleasure produced by my actions? Are other considerations important, such as the quality?

**Activity**



Consider the following scenario:

*Dr Achebe is a keen angler and has decided to go on a fishing holiday in the Caribbean. On the third day of his holiday, he joins an expedition with four others to a remote island where the catch is said to be particularly good. En route, a tropical storm of fearsome magnitude envelops the small fishing boat and the party find themselves shipwrecked. The four other anglers, all of whom have sustained serious injuries, are left floating in a small life raft, unsure if or when they will be rescued. Fortunately, Dr Achebe has packed some medical supplies; unfortunately, he only has enough to treat half the group. Therefore, he must risk letting two people perish.*

Over the course of the trip, in conversation Dr Achebe has learnt that:

• Fisherman A is a 57-year-old South American priest who has done much charitable work in   
his community.

• Fisherman B is a 43-year-old, who works in the accounting department of a large tobacco company. He is happily married with five children.

• Fisherman C is a 27-year-old bachelor who writes screenplays for big-budget Hollywood films.

• Fisherman D is a 32-year-old widower who works part-time as a handyman. He has one teenage son.

Using the utility calculus, try to work out which two people Dr Achebe should save. Afterwards, discuss your results with the rest of the class. Did everyone come to the same conclusion? Does this experiment raise any issues for Bentham’s theory?

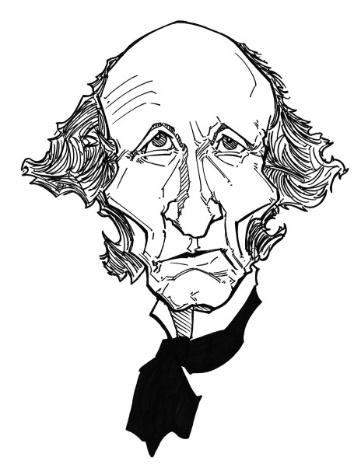
#### Bentham, Motivation and Consequentialism

Bentham’s utility calculus demonstrates well how utilitarianism is a consequentialist theory. When considering the right action, for Bentham it is only the consequences that matter; the most amount of pleasure produced by each action, and the intention of the agent does not factor into whether an action is considered right or wrong. Many people regard this as intuitive – we often like to believe our good intentions should be rewarded, even if the action we perform doesn’t necessarily have good consequences.

One instinctive difficulty nevertheless arises when we consider what should be our motive for moral action. For Bentham, the primary motivation for human beings is pleasure, yet at the same time the right action is one that produces the most amount of pleasure for everyone involved. While this might be explained by a general altruistic intention people might possess, Bentham does not highlight this as an important theory of his moral theory. Rather he presents a strong legalistic emphasis in his utilitarianism – it is the role of government to ensure laws are created and punishments enforced that grant pleasure for being socially responsible, and pain for when people focus on only their pleasure. Therefore, the person who takes pleasure in smashing postboxes at the weekend is punished such that their motivation to engage in the postbox smashing is muted by the pain it entails by being jailed or fined. Similarly, the person who pays their taxes and follows their duties is given wider freedom to engage in activities that grant them pleasure.

Yet many find this tendency of Bentham to view people as pain or pleasure machines uncomfortable to stomach. It naturally implies governments have a responsibility to manipulate humans and force them to conform to a correct maximisation of pleasure over pain, a view that many people would oppose as social engineering. If one has to resort to subtly influencing human beings to follow certain cues of pleasure and pain one might argue one has not really morally educated people, only trained them as a person would an animal.

Mill’s Utilitarianism

John Stuart Mill (1806–73) is still perhaps one of the most prominent and influential liberal philosophers in history, contributing not only to an expanded and nuanced theory of utilitarianism, but also to political theory, economics and social theory. Both his 1859 text *On Liberty* and the 1863 *Utilitarianism* have greatly influenced discussion around moral philosophy and its relation to the creation of the ideal state, from which human beings can flourish and develop. Throughout this section, we will discuss the different ways Mill sought to improve Bentham’s thought, and to capture important elements of moral thought that many perceived Bentham to have overlooked.

#### Mill’s Proof of Utilitarianism

Mill fundamentally agreed with Bentham that the principle of utility was the foundation of morality. As he describes it:

Utility, or the Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is intended pleasure and the absence of pain; by unhappiness is intended pain, and the privation of pleasure.

**Anthology Text:**

John Stuart Mill (1879), *Utilitarianism* Ch. 1, 2, 4, 5



However, Mill, more so than Bentham, believed that the principle of utility could not be assumed to be true from a simple analysis of human nature; it required a reasoned proof that shows there is a strong foundation to utilitarian thought in the principle of utility. Yet Mill did not view his argument as a formal proof, but rather a reasonable assertion based upon certain agreeable facts about the world and human life.

As such, the argument does not take up a huge amount of space in utilitarianism, and critics of Mill have often focused on what they view as the fallacious foundations of Mill’s argument, in particular its susceptibility to the **naturalistic fallacy**, and the **fallacy of composition** (which will be explored later). Nonetheless, many contend that such criticisms focus on a misleading assessment of Mill’s proof, and, examined within his wider empiricist beliefs, his proof stands as a reasonable foundation for utilitarianism.

Mill begins Chapter 4 of *Utilitarianism* by asserting what he believes to be a simple truth – happiness is desirable. How does he arrive at this basic truth? Firstly, by asserting a simpler idea, namely that each person desires their own happiness:

The only proof capable of being given that an object is visible, is that people actually see it. The only proof that a sound is audible, is that people hear it: and so of the other sources of our experience. In like manner, I apprehend, the sole evidence it is possible to produce that anything is desirable, is that people do actually desire it … No reason can be given why the general happiness is desirable, except that each person, so far as he believes it to be attainable, desires his own happiness.

Here, Mill is making a basic empirical claim; people desire their own happiness and the only thing that justifies such an assertion is the observation that people in fact in their lives do desire their own happiness. It might be possible to disagree with Mill here, but one would have to argue such a claim is wrong and people in fact do not desire their own happiness, which in many ways is a more controversial claim than Mill’s.

This leads to Mill’s second set of claims; that since people desire their own happiness, happiness for each person is a good to them, and so general happiness is a good to all people.

… happiness is a good … each person’s happiness is a good to that person, and the general happiness, therefore, a good to the aggregate of all persons. Happiness has made out its title as one of the ends of conduct, and consequently one of the criteria of morality.

Many might point out a flaw in reasoning, that simply because happiness is desired for each person does not mean that general happiness is desired for all people, and this potential fallacy will be explored in the next section. Yet Mill does not view this jump as that contentious. If one can accept that happiness is good for one person, and is good for all others, one can arrive at the conclusion that general happiness is a good to all people as an aggregate.

The final part of Mill’s proof is perhaps the most contentious claim he makes. While his proof until now has identified happiness as ‘one of the ends of conduct’, it has not established it as the only end. In fact, as Mill notes, there are many other possible ends, one of the most prominent being virtue. Therefore, there is potential competition for happiness. While people seem to desire it, it may be the case that people really desire virtue, and happiness is a product of this virtue. Similarly, he notes people may regard money as the end of all action, with that being a primary desire of human beings beyond happiness.

However, Mill argues that all these other ends, whether they be virtue, money or music, are originally pursued because they produce happiness for a particular person, and eventually through this continual association with happiness they appear to become ends in themselves. This means that ends such as virtue, from their origins in happiness, should really be seen as supporting happiness being the sole end of human action. In fact, Mill argues that it is an important feature of human existence that there are things which produce more happiness than the ‘primitive pleasures’ that would ordinarily satisfy human beings. Without some greater concepts governing our lives than the simple pursuit of base pleasures, life would be much poorer in quality.

Mill’s ideas of psychological association were common during his era, and were often used to explain numerous philosophical and conceptual phenomena, and how they arise from ordinary experience. The importance given to virtue by many human beings as such is simply a product of a continual association with a certain state of pleasure. As Mill describes:

Virtue, according to the utilitarian conception, is a good of this description. There was no original desire of it, or motive to it, save its conduciveness to pleasure, and especially to protection from pain. But through the association thus formed, it may be felt a good in itself, and desired as such with as great intensity as any other good…

Many people have criticised Mill here on both psychological and philosophical grounds, which will be explored in more detail later. But it is important to note that this part of Mill’s argument is perhaps the most difficult to defend. Virtue may well not be easily associated just with states of pleasure, and there may be pertinent grounds to argue that happiness is not the only end humans seek, but is one of many, or a by-product of more important ends.

#### Discussion:

*Do you think accounts of psychological association are convincing? What could potentially prove Mill   
wrong about his ideas concerning pleasure?*

#### Is Mill’s Proof Fallacious?

What is a fallacy? Many different answers have been given to this question, but the simplest way of defining a fallacy is as an error in reasoning that causes a particular argument to be weaker than previously thought. These can be separated into ‘formal’ and ‘informal’ fallacies. The former are fallacies that occur in a fault in the form of an argument, and are often seen to be the most damaging form of fallacy, as they usually mean that the conclusion one draws from an argument is not supported by the premises. An example of a formal fallacy would be:

If I drink hot chocolate, it will snow outside.  
It is snowing outside.  
Therefore, I drank hot chocolate.

Why is this a formal fallacy? Well, the argument puts forward that when I drink hot chocolate, it will snow, but this doesn’t mean that it snowing is always dependent on myself drinking hot chocolate. It might be snowing for a different reason; therefore, it cannot be concluded that every time it snows I must have been drinking hot chocolate, only that when I do drink hot chocolate, it will snow. The structure of the argument is incorrect, even if the premises prove to be true, making it a formal fallacy.

However, informal fallacies are different; they require an analysis of the content of an argument, not its structure, and as such they are often more difficult to refute an argument with, since they often involve discussion of what different words and terms mean, which people often disagree on. It is these forms of fallacies that Mill’s proof has been argued to fall prey to, and two important ones will be examined below.

#### Fallacies

Criticism in philosophy is more often than not focused around finding the hidden fallacies in arguments   
that initially seem sound. Very rarely is an argument accepted without modification or critical analysis.

You can find a list of informal fallacies here: *https://en.wikipedia.org/wiki/List\_of\_fallacies*

Pick three and think about where you might have seen them occur in real-life scenarios. A key part of developing skills in philosophy is learning how, when and why fallacies might be present in different people’s arguments and to what extent they are revealing of the weaknesses of a particular argument.



**Stretch and Challenge**

Go to **zzed.uk/9765**

🖰

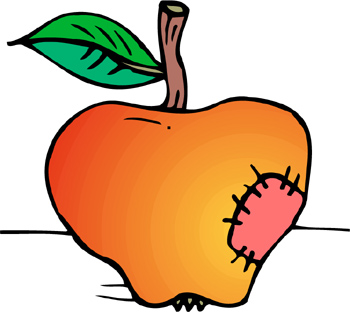
#### The Fallacy of Equivocation

The first is the **fallacy of equivocation**. This is where a person defines a particular term or word in an argument as having one meaning, when in fact it may have multiple meanings. The fallacy of equivocation is often common in ethics, as there are many different moral terms human beings use that don’t have a fixed meaning, and this is where the philosopher G E Moore criticises Mill’s argument.

It was noted how Mill argues that all people desire their own happiness, and this leads to the principle of utility; that this happiness ought to be maximised. Yet Moore argues there is a mistake in reasoning here. Mill assumes that ‘what people desire’ is the same as ‘what people ought to desire’. Why is this important to point out? Well, it potentially assumes too much in the term ‘desire’, for there are potentially many things that people desire that many people would argue ought not to be desired. For example, I might desire before work in the mornings to throw rotten fruit at local church windows, but such a desire might not be what I ought to desire. It might have been born out of frustration at my boredom sitting through church services at a young age.

Therefore, there seem to be two different ways one can understand desire. The first is purely descriptive, or factual, pointing out what people do desire, or can desire. The second is a moral understanding of desire, which covers things and states that ought to be desired. In the case of Mill’s proof, Moore argues that Mill assumes that one can reasonably move from what people do desire into an argument about what ought to be desired, and that this assumption is unwarranted.

#### The Is–Ought Gap, and the Naturalistic Fallacy

The fallacy of equivocation in Mill’s proof highlights an issue in ethics that has been prominent since the work of David Hume in his *A Treatise of Human Nature*, written in 1739. This lays out what has come to be known as the **is–ought gap** in ethics; the idea that you cannot logically derive what ought to be the case from what is factually the case. Therefore, no matter how hard I try, I cannot logically justify that people ought to throw rotten fruit at churches, simply because I choose to do for some factual or descriptive reason.

This gap will be explored more in the Meta-ethics section, but it led Moore to identify what he termed the naturalistic fallacy in Mill’s proof. This puts forward that any move to identify what is morally ‘good’ through equating it with some natural fact or property is fallacious, for if we define something natural as good, it is still possible to ask whether that natural thing is in fact good. If such a question is meaningful, then it cannot be the case that the natural thing and good are the same. For example, if I defined what is good as that which gives me the most money, I could still ask ‘but is that which gives me the most money actually good?’ If that which gave me the most money was actually good, then that question is really meaningless, as it could be phrased as ‘is that which gives me the most money actually that which gives me the most money?’ What this naturalistic fallacy highlights for Moore is that there is something unique in the moral way we define what is ‘good’. It is not simply something natural, such as happiness, and as such it is important to note that one can argue against utilitarianism by arguing that its definition of the good isn’t correct.

#### The Fallacy of Composition

Another informal fallacy Mill makes is the fallacy of composition, which is where one assumes that what is true about part of something must be true about the whole thing. For example, simply because a house is made out of bricks does not mean that the house itself is a brick. The fallacy of composition, therefore, trades in what words mean in different contexts and what we consider to be the properties of individual things versus the properties of the things they form as a whole.

Where does Mill make this fallacy? Well he arguably assumes that from identifying happiness as a good to each person, one can say that the general happiness is a good to the aggregate of peoples. Yet this isn’t necessarily true, and it may be the case when analysing a general aggregate of peoples that we can’t ascribe desire in the same way we can to individual people. Rather different terms and ideas may apply, of which happiness is not a valid one.

#### Are Utilitarian Ethical Theories Unjustifiable?

Nevertheless, the greater question beyond identifying these fallacies is whether they really weaken Mill’s proof of utilitarianism. As informal fallacies, we have seen how they rest on interpretations of content, and as such they can be contested by arguing that Mill did not make a mistake in interpretation in his proof.

For example, regarding the fallacy of composition, it can be argued that Mill is not setting up a logical proof regarding general happiness, rather an empirical claim; that if we accept each of us desires happiness, it is evidence that from an impartial perspective the general happiness can be thought to be something to be desired. One might dispute the jump, but evidence that Mill is wrong and we don’t wish the general happiness of all would have to be provided.

Where Mill perhaps faces the greatest difficulties are the naturalistic fallacy, and his theory of association regarding the ends of human conduct. The former fallacy, however, affects all naturalistic ethical theories, as will be examined in the Meta-ethics section, and once again it can be stated as a matter of empiricism that if morality does not come from experience of the world, where else might it arise? Such questions are difficult to answer. Similarly, if other ends of human conduct such as virtue, or aesthetics, do not arise from an original association with pleasure, where else might they have arisen, and what sort of status should we give them in relation to happiness? Therefore, while one can identify informal fallacies in Mill’s argument, there are not conclusive faults and the utilitarian still has strong grounds for their form of normative ethics.

### *Higher/Lower Pleasures*

Bentham was Mill’s godfather although the two men had rather different characters. Mill’s father, James Mill, was himself a utilitarian philosopher and economist who was keen that his son be capable of carrying on his intellectual legacy. As a result, Mill’s childhood was devoted almost entirely to learning; he was studying Greek at age 3, Latin at age 8, and by his mid teens was well acquainted with the works of Plato, Aristotle, the classical Greek poets, and also a number of Victorian economists. In his spare time, he had also managed to develop a solid grounding in higher mathematics, logic and the natural sciences. He was, however, to suffer a mental breakdown in his early twenties, which he attributed to the abnormally rigorous educational regime he was subjected to by his father. Mill was said to have only recovered with the help of Wordsworth’s romantic poetry.

Mill’s sensibilities are reflected in his more nuanced account of pleasure. For Bentham, all pleasures were in a sense equal; it was simply a case of quantities. Mill, on the other hand, is much more concerned with the *quality* of pleasure. His thought was that there is something different about pleasures of the mind such as art, literature and philosophy and bodily pleasures such as sex or drink. It’s the difference between a fine cut of beef expertly prepared in a Michelin-starred restaurant, and a 99p cheeseburger from a disreputable burger van.

Mill states the test for determining whether a pleasure is of a higher quality than another as follows:

Pleasure P1 is more desirable than pleasure P2 if: all or almost all people who have had experience of both give a decided preference to P1, irrespective of any feeling that they ought to prefer it. (*Utilitarianism*, Ch. 2)

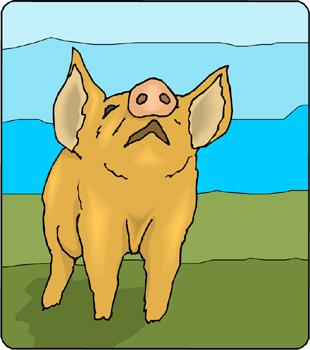
**Group Activity:**

In groups or as a class, make a list of the activities you find pleasurable. Then apply the test above. Which are the higher pleasures, and which the lower?

However…

Is it not a little idealistic to suppose that people will always choose, for example, going to the opera over a bucket of chicken and *Coronation Street*? Indeed, might it be the case that often the so-called lower pleasures are far easier to satisfy (in terms of both availability and effort) than the higher ones? For example, appreciating a dense work of high modernism such as James Joyce’s experimental novel *Ulysses* requires not only a significant investment of time but also considerable intellectual resolve (it’s not an easy read!). Is it not far less effort and far quicker to just watch an Adam Sandler film for a few cheap laughs?

Bentham would agree, for he saw all pleasures as equal, whether it be reading *Ulysses* or watching Adam Sandler. One can question whether the act of differentiating pleasures in this way is simply a case of intellectual snobbery. For example, it is often the case that those with the greatest sensibilities, intellectual refinement and so on are the most likely to succumb to melancholy (as Mill himself did). Is it not in fact better to have only those desires which are most easily fulfilled? Why torture yourself for art or the intellect when you can have an easier life just getting drunk and watching television?

In a famous passage, Mill responds to this objection by arguing that there is a distinction between *happiness* and *contentment*. Those who are better able to use the ‘higher faculties’, Mill claims, may be less content but they are still happier. This is because they know of a *greater* happiness which is *unavailable* to those who are satisfied only by the lower pleasures. Those who have access to both kinds of pleasure, know immediately how much finer it is. They are the cat who does not want to go back to milk now it has tasted cream. As Mill puts it: ‘*It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied. And if the fool, or the pig, are of a different opinion, it is because they only know their own side of the question.’* (*Utilitarianism*, Ch. 2, p. 7)

J J C Smart, a modern utilitarian, makes a similar point. One could have their base desires satisfied and live in contentment, but this does not cover what happiness really means. Rather there is a secondary component which regards happiness as also where we are positively satisfied with how we come to be contented. These can be thought of as ‘second-order’ desires that influence how we wish to be contented. For example I might be satiated by watching my Adam Sandler film collection, but am I really satisfied with a life only gaining pleasure by rewatching *Happy Gilmore*? I might think at the back of my mind that there must be more to life, and wish for experiences of a greater quality, even if they aren’t as immediately rewarding.

#### The Competent Judge

But how to know which experiences are qualitatively more rewarding? Mill argues this question can only be decided by those who have had an appreciation of the experiences of concern and can choose which are higher and which are lower. Therefore, whether Adam Sandler films are a lower or higher pleasure compared to *Ulysses* can only be decided by people who had consumed both and are in a position to evaluate their quality. These people can be regarded as ‘competent judges’, in the same way we might defer our opinion of certain forms of art to qualified critics.

Mill argued this kept the distinction between higher and lower pleasures reasonably objective. Although Mill recognised there would be disagreement, he still believed that his position on the greater forms of pleasure, such as art, music and philosophy, would be validated by said judges. In the end, however, under Mill’s system, if the competent judges decided Adam Sandler was the higher pleasure, such a decision for Mill would have ‘to be admitted as final’. Whether or not such a conclusion would ever be reached is another matter entirely!

#### Discussion:

*How do we judge the competent judges?*

*Say, for example, we have two individuals, both versed in Adam Sandler and James Joyce, having consumed all their works numerous times. However, at the end of this process, they both fundamentally disagree about which are higher and lower pleasures. Would such disagreement be a difficulty for Mill’s philosophy?*

#### Is Mill’s Utilitarianism Truly Hedonistic?

It has been asked, if we are invoking some other principle to discern which pleasures are more important than others, is Mill’s utilitarianism really hedonistic? Some have argued real hedonism does not involve such qualitative measures of pleasure, rather a simple aim to maximise whatever pleasure is deemed available to individuals.

However, Mill does not have to be seen to be making a radical claim about pleasures. When we compare two activities, we don’t just analyse how long and intense the pleasure was, but how that particular pleasure compares in its nature and quality to others. Human beings naturally assess some activities as being more important than others, and this helps them work out where to focus their energies. Yet utilitarianism does not just discuss pleasure on a personal level, but also on a general level; hence the need for objectivity in the competent judge. As Mill states:

I regard utility as the ultimate appeal on all ethical questions; but it must be utility in the largest sense, grounded on the permanent interests of man as a progressive being.

Critics might still argue that when we talk about competent judges, we really are invoking a new secondary principle to evaluate hedonism, implying something extra is contained in happiness beyond pleasure. However, it is important to note that Mill simply regards his point about higher and lower pleasures as one identifying what is necessary in a sophisticated account of human happiness, not just equating happiness with raw pleasure.

#### The Disadvantages of Mill’s Utilitarianism

A final word can be said about the higher/lower pleasure distinction, namely that although it seems to present a more sophisticated account of happiness, it also greatly complicates the utilitarian project. Cast your mind back towards Bentham’s utility calculus; it was already noted how complex evaluating pleasurable outcomes was when we were only discussing the quantity of pleasure; if certain pleasures are incommensurable (unable to be directly compared with the same standards), how on earth could it be possible to discern the correct moral outcome of an action? For example, how many Adam Sandler films would I have to watch to reach the same pleasure as reading *Ulysses*, or are the works of Joyce always given a priority to the back catalogue of Sandler’s films irrespective of how many new comedy films he produces? Should we only show French art-house films instead of lowbrow comedies in cinemas?

Before you argue such questions are arcane, and not relevant to ordinary life, note that such considerations pertinently do have real-world impacts. If a government official has to decide on whether to fund a new football stadium for 100,000 West Ham fans, or a new opera house for a select few Handel enthusiasts, Mill’s standards may mean they should always choose the latter, regardless of however unjust a decision it may seem. The unintended consequences of Mill’s theory may be a persistent intellectual snobbery that results in a detrimental impact on those who truly love football (and Adam Sandler’s films).

Overall, therefore, the simplicity and ease of use that is supposed to be an advantage of utilitarian theories is weakened, and, although some issues are answered in Mill’s theories, many more are arguably brought up in the process. How can these be addressed? One possible solution is presented by Mill in the next section.

Act and Rule Utilitarianism

Another aspect of Bentham’s philosophy that Mill rejected outright was the former’s ‘utility calculus’. Mill argued that happiness was ‘much too complex and indefinite’ to be calculated in every ethical situation. What did Mill mean by this? Well, noting the difficulty we had in judging basic decisions about whether to watch Adam Sandler films or read James Joyce, it can be noted the number of ethical decisions one might have under utilitarian theories is potentially endless. When is it exactly that we are required to think about the outcomes of our actions, and calculate the correct moral response? Is it every time we are presented with different possibilities and outcomes, or only when such outcomes might have a significant impact on ourselves or others?

#### The Implications of Act Utilitarianism

**Happiness or pleasure?** Both Mill and Bentham equated happiness with pleasure; Mill, however, measures pleasure qualitatively, whereas Bentham measures it quantitatively. Both also argued that the only good in life *is* pleasure, a position known as **hedonism**. Don’t be surprised, then, to see Mill refer to the principle of utility as ‘the Greatest Happiness’ principle; they amount to the same thing.

Such questions are ones which the act utilitarian may find it difficult to answer, and, despite its simplicity in approach, means it is complicated to put into practice. For such reasons many have argued that act utilitarianism is simply too burdensome on the moral agent. Either it is impossible to work out on a consistent basis, or it fails to apply itself consistently when we decide to evaluate moral outcomes in one situation and not another.

However, a further criticism can be made of act utilitarianism. Are all acts morally justified so long as they produce overall greater happiness or pleasure? Consider the following examples:

A) Arthur is mayor of a town set next to a giant mountain, where in a cave resides a powerful dragon. Every year the dragon comes down to the town and demands that one child be provided as a sacrifice in order to keep it satisfied. If not it will burn all the houses in the town to the ground. The mayor, thinking that the disappearance of all buildings in the town will cause grave unhappiness to everyone, not just the family of one child, agrees each year to sacrifice one child to the dragon, preventing widespread destruction and pain.

B) Terry becomes ill one day, and, realising he might have appendicitis, goes into hospital for an operation. However, the surgeon realises Terry has two well-functioning kidneys, and, due to a shortage of overall kidneys for patients needing a transplant, secretly removes one of Terry’s kidneys, leaving no trace of his action. Terry goes on living the rest of his life healthily, unaware he only has one kidney, and another person gets a new kidney, saving their life.

C) Betty, the head of a local psychiatric hospital, realises she has a problem. There are a number of patients who openly wish to inflict harm and kill other people, and, being denied this pleasure, are deeply depressed and often tempted to harm staff and other recovering patients. However, she also has a number of severely depressed patients, who regularly attempt suicide and often self-harm for pleasure. She decides that the best solution is to let the violent inmates occasionally, under a controlled environment, kill the depressed patients. Such a decision seems to bring all parties great pleasure, especially those violent inmates, who become much happier and no longer attack staff and recovering patients on a regular basis.

What is potentially wrong with the moral decisions each person makes in these situations? They promote what is arguably the outcome with the greatest overall pleasure, yet it is likely that at least one or two of the examples go against your moral intuitions. A number of reasons have been given for such a contradiction. The first and most prominent, which will be explored later, is that act utilitarianism gives no guarantee of rights, for no action is off limits for any reason other than it does not give the greatest amount of pleasure.

However, there are other potential issues. Consider example A; it seems that no matter what Arthur does, there is inevitable misery involved for the town. Yet Arthur has decided that even if the sacrifice of the child is only marginally more pleasurable for everyone than the burning of the entire town, it is still the right action. Many would argue in such cases greater principles, such as justice or righteousness, should be considered, and, in order to ensure no more pain in the future, even if it means obliteration or deep permanent pain, fighting the dragon and preventing the sacrifice of any more children.

Similarly, we can as such imagine cases where, faced with an ethical dilemma, one might be faced with a morally neutral action, or one that causes both great pain and a marginally greater amount of pleasure. Under act utilitarianism, no matter how much pain is caused we are always required to pick the more pleasurable action, and so, even in ordinary circumstances, we might be inclined towards extreme acts all the time simply to generate an eventual slight increase in overall pleasure.

What these circumstances indicate is that human beings do not just consider what is pleasurable when deciding upon moral action. Other principles come into play, such as our understanding of justice, or moderation, or even natural freedom. So how are these principles captured under a utilitarian system?

#### Rule Utilitarianism

Mill’s solution was the introduction of what he termed ‘secondary principles’ and, as such, outlining what can be called **rule utilitarianism**. This puts forward that for ease of action and consistency, people should follow certain rules and principles that lead to the greatest maximisation of pleasure for all. Therefore, even if one is faced with a situation where an action seems to break a rule and bring greater pleasure, one may be justified in following the rule, even if it results in greater pain. Therefore, human rights can arguably be established, along with utilitarian rules that emphasise the importance of justice or virtue, in the belief that following such principles at various times is more likely to bring greater happiness than only focusing on what is pleasurable.

**Rule Utilitarianism**

The theory which holds that the right action is one which follows rules that, if universally obeyed, would create the maximum amount of pleasure (or happiness) and the minimal   
amount of pain.



However, how should we consider these rules in relation to the core utilitarian aim of maximising pleasure? While inevitably they make utilitarianism easier to follow, what do we do when following a particular utilitarian rule seems transparently wrong? Isn’t an advantage of utilitarianism also its flexibility to make the right moral choice against what might seem initially morally intuitive?

We can, therefore, distinguish a ‘strong’ and a ‘weak’ rule utilitarianism:

**Strong Rule Utilitarianism:** One ought to always obey utilitarian rules, even if breaking one seems to bring significantly more pleasure.

**Weak Rule Utilitarianism:** One ought to always obey utilitarian rules, except in exceptional circumstances where breaking a rule results in a significant maximisation of happiness.

#### The Issue of ‘Rule Worship’

It might be asked, what is the point in creating rules if they will inevitably be broken when evaluated as wrong? Rules aren’t really rules unless they are binding.

Yet such a question ignores the weak rule utilitarian’s commitment to the primary rule; that one ought to maximise happiness. When one breaks the established rules it is not on an arbitrary basis, but when the established rules conflict with themselves, or with the primary duty to maximise happiness. This is what Mill envisaged when he outlined the place of his ‘secondary principles’:

We must remember that only in these cases of conflict between secondary principles is it requisite that first principles should be appealed to.

Such a distinction avoids what J J C Smart calls ‘rule worship’, the counterintuitive circumstances of a utilitarian disobeying the primary principle of utility simply to follow rules that are supposed to maximise utility. Yet even weak rule utilitarianism has its own share of problems, which will be analysed in the Issues section.

We can ask at this point, what status do these secondary principles really have? If we are still analysing everything really in terms of utility, have we properly created real principles of justice and rights, or are we just simulating them to avoid the philosophical dilemmas of act utilitarianism? Many would argue that there is an unavoidable fork for the rule utilitarian; either justifying rule worship against the principle of utility, or simply endorsing a more complicated and unworkable version of act utilitarianism.

The Strengths of Hedonistic Utilitarianism

Before we analyse the deeper issues in Bentham’s and Mill’s utilitarianism, it is worth considering why, on the face of it, utilitarianism is an attractive system of normative ethics. For both Bentham and Mill were aware of its shortcomings at the time of writing, yet continued to believe it set apart the right way forward when discussing ethical dilemmas. Why?

 **Simplicity** – Although we examined how it might struggle to solve different moral dilemmas, this is arguably an issue for all normative ethical theories. If ethics were easy we wouldn’t have such dilemmas in the first place! But underlying utilitarianism more than any other normative ethical system is a simple assertion: we should seek to maximise happiness, and its one that at first glance is uncontroversial. We all want to be happy generally, and we generally want others to be happy too.

 **Flexibility –** At the same time utilitarianism is a theory that is able to accommodate a wide number of factors and individuals when making a moral decision. We are not excluding facts about the world or people when discussing what is moral, and we don’t ignore key elements of our decisions that might override the happiness of others. No matter what situation we are in, we can evaluate to a certain degree what might result in the greatest happiness, and give reasons to support that decision.

 **Applicability** – There is arguably never a moment in our lives when we do not consider our happiness in some way. Many people would argue this process of thought transcends cultural boundaries. While two societies might have different understandings of justice, or righteousness, human beings naturally understand pleasure and pain, and can potentially understanding the importance in maximising these aspects of their lives.

 **Effective when Lawmaking** – We’ve seen how Bentham developed utilitarianism to help organise the legal system, and it certainly is an advantage of utilitarianism that it helps evaluate how we might develop laws, rewards and punishments in a society. It can measure the impact of actions we believe bring unhappiness, and legislate based upon the harm people inflict on others, with recourse to more vague moral concepts and ideas.

 **Impartiality** – This is a key advantage for utilitarianism in that it does not generally grant some people’s desires and wishes more importance than others. All people are assessed as producing happiness or unhappiness, and within the moral system there is no difference between rulers and ordinary workers when assessing moral action. To some, utilitarianism represents a truly impartial ethical system that guarantees equality between people when considering preferences and desires.

However, these initial strengths can be questioned and criticised, and in the next section it will become apparent that the case for utilitarianism is not so clear-cut.

Issues for Hedonistic Utilitarianism

1. Individual Liberty and Rights

#### Bentham on Individual Liberty and Rights

Bentham notoriously claimed that the idea of natural rights is ‘nonsense on stilts’; his thought was that since something can only be called a right if it is respected as such, then it makes no sense to talk about rights people have naturally, because the only way to have a right is to have a right *by law* (a social institution). Contrast this view with this line from the United States Declaration of Independence: ‘We hold these truths to be self-evident, that all men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty and the pursuit of Happiness.’

The practical consequences of this position are that human rights, as we may conceive them today, do not exist. Most people would be unhappy with such a conclusion, including Bentham and Mill! For it allows tyrannical rulers, or even just ordinary people, to use others so long as it contributes to overall happiness. We might imagine a society where one third of the population is kept in abject poverty and slavery in order to finance the luxurious lifestyles of the other two thirds.

There are innumerable scenarios that make the act utilitarian views of Bentham difficult to square with how we might wish to mould society and establish rights (see pp. 14–15) but such considerations seem to require fairness as a principle in ethics, and many argue that act utilitarianism cannot simply put forward a consistent principle that would make our moral interactions fair between peoples.

#### Mill on Individual Liberty and Rights

Mill was aware of such issues, and wished to develop his utilitarianism partly to avoid a ‘tyranny of the majority’ where the rights and voices of minorities would be protected, so that not only a self-interested majority could benefit at the expense of everyone else. This led to his emphasis on the importance of liberty and the outlining of the ‘harm principle’; that one is permitted to act as one wishes except where it harms others. This can be seen as the basis for a more extensive system of rights, for it does not allow others to inflict pain on me simply to gain a greater pleasure for themselves.

Yet even under Mill’s weak utilitarianism, we can question whether we really have true human rights, for we have already identified that we are required to often resort to the principle of utility to resolve conflicts. This means that political leaders may still be allowed to violate human rights in what they consider exceptional circumstances, and many people would argue even this situation is unacceptable. Human rights, for them, should not be violated, and so long as a leader possesses the theoretical right to violate rights then one is subject to arbitrary domination, meaning one is not really free, and does not really have rights in the important sense of the term.

What this means is that it is still far from clear whether Mill can justify rights under his form of utilitarianism. Rather it may be the case that human beings don’t have guaranteed rights, only conditional rights based on certain factors. Some individuals may be happy with such a conclusion, acknowledging sometimes the greater good is more important, even if it impacts the individual. However, others may be more critical, and it is important to note that whether utilitarianism can justify real rights is still a matter of debate to this day.

***Mill’s Political Philosophy***

It is also worth noting that Mill’s other famous work of philosophy, *On Liberty*, is regarded   
as one of the seminal texts of political liberalism. At the heart of the book is the harm principle:   
‘The only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others.’ In effect, this means that the state should be a ‘neutral umpire’ which upholds individual freedoms and only interferes in society to prevent individuals from harming one another. Research Mill’s arguments in *On Liberty*, and consider whether his ethics consistently supports his political ideals.



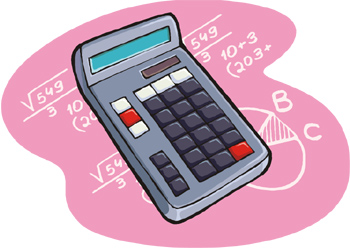
**Stretch and Challenge**

2. Problems with Calculation

#### Bentham on Calculation

Bentham may have thought he had solved the problem of how pleasure is calculated by introducing the utility calculus, but arguably he has only dug himself into a deeper hole. Firstly, time is rarely in abundance when an ethical decision needs to be made, making Bentham’s multifaceted utility calculus seem almost laughably impractical. For instance, would the narrator of the *Trolley Problem* have time to whip out pen, paper and a calculator in order to decide whether or not to push the button and switch the tracks?

This addresses an important issue in ethics: in what way should we hold individuals morally responsible for their actions? Utilitarianism would seem to require we hold individuals responsible based on the consequences of their actions, but what consequences should people reasonably foresee, and should people be held responsible for consequences they don’t foresee?

However, the criticism is perhaps unfair. What consequentialists expect us to base our decision on is *our best guess* of what is going to happen; they do not expect absolute precision because life is simply too unpredictable. This is a reasonable requirement; there is, after all, much about human behaviour that can be predicted. Were a government agent to follow me for long enough, they would soon be able to predict where I would go after work on a Tuesday night, who I might meet there; if they had been listening in to my calls, they might even be able to predict the sorts of things I would discuss. (We will come back to this point again when we look at the issues with Kant’s deontology.)

Nevertheless, even making an educated guess about the outcome of a particular action may be asking too much. The issue Bentham’s theory faces is that it is overly complex, a difficulty Mill was keen to resolve.

#### Calculations and Right Action

Another issue that can be explored is that intentions for the utilitarian don’t seem to matter. So when we are calculating right action, it seems as if goodwill shouldn’t play a part. Is this really how human beings think? Moreover, how do we differentiate in responsibility between the more and less intelligent people? For example, should Einstein be held to a higher degree of moral responsibility than a person who is intellectually impaired? What counts as a best guess of what is going to happen is difficult to discern, and it seems as if a good moral theory should at least be able to capture an instinctual element of how human beings ethically think, such that in a dilemma, both the more and less intelligent people can at least discern the beginnings of right action. For a greater discussion of this, see Issue 4: Character and Motives.

#### Average vs. Total Happiness

Imagine the following scenario: in 2050 there are now two worlds on two different planets in the solar system, and a group of utilitarian philosophers have been tasked with deciding which the best is, so that the less-well-off planet can be remodelled in the other’s image.

World A is set on Earth, and is now very overpopulated, with 10 billion people alive at any given time. This has resulted in declining living standards for many people with ever increasing pollution and resource shortages. However, people are making do, and the utilitarian philosophers viewing a range of happy attitudes from 1–5 decide they can roughly give a net worth value of 3 for each person’s happiness.

World B is a new colony on Mars, with a population of 2 billion where the planet is terraformed and resources are still in great abundance. People are living in comparative luxury and so the utilitarian philosophers decide they can give the highest net worth value score of 10 for each person’s happiness.

On the face of it, we all would intuitively argue that World B is the better world. People are generally happier and the environment is better managed. However, this presents a difficult problem for the utilitarian, for the principle of utility says to maximise happiness, but not whether this means the average or total happiness. If we take total happiness to be important, then World A is superior, and World B should reproduce until they reach the population of World A, even if this reduces the average happiness for everyone. However, if we argue that average happiness is important, then it would seem that we should kill or get rid of anyone on World A who has a lower than 5 happiness value to improve the average to the highest it can be. In fact it can be contended that the logical conclusion of maximising average happiness is to kill or get rid of all but the happiest person, thereby ensuring the highest possible average happiness score.

The utilitarian might retort that such examples are too hypothetical, and do not really apply to the real world, where a variety of factors and influences affect how each person relies on others for their happiness. Yet this does not prevent this question from being an issue for utilitarians, and how one judges how to measure happiness potentially influences a vast array of modern political and social issues, including healthcare, population management, inequality and social justice.

#### Discussion:

*Wealth inequality is thought to be a growing problem around the world. The latest World Inequality   
Report calculated that between 1980 and 2016, the richest 1% of the population captured around 27% of the total wealth of the world, roughly the same as the bottom 50%. (*http://wir2018.wid.world/*)*

Go to **zzed.uk/9765**

🖰

*How might a utilitarian approach this issue in terms of calculating average vs total happiness? Do you think   
both approaches would yield the same conclusions, or do you think one would be preferable over the other?*

#### When Does One Finish Calculation?

One of the difficulties consequentialist ethics often face is deciding at which point one should end moral calculation. For example, say one day you are crossing the road and just in front of you, you see an unaware child about to pass into the path of a car speeding through a red light. Instinctively you drag the child backwards, saving their life. However, 20 years on, as you relax in your living room watching television, a sudden news bulletin appears on the screen informing you that the same child you rescued is now personally responsible for a severe environmental disaster, costing the lives of thousands of people through extremely unethical business practices.

Looking back you might wonder whether you were right to save the child. In fact a real-life case of this potentially exists (although many believe it is an urban legend). A well-known soldier, Henry Tandey, claimed that in the First World War he spared the life of a man, who later turned out to be Adolf Hitler. Under utilitarian ethics, while this decision may have initially seemed to be a positive act of mercy, it should instead perhaps be regarded as one of the worst moral decisions in the twentieth century.

Many utilitarians might argue that such questions are avoiding a simple solution; that people should only be expected to reasonably calculate the foreseeable outcomes of their action. One might, therefore, make the distinction between outcomes one should be required to foresee, and those which lie outside any boundaries of foresight. This might work on small scale, everyday decisions, but what about those in which the moral agent understands the potential long-term impact of their decisions? A person planning the building of power plants might have to weigh up the potential impacts of blackouts, and the civil unrest it may cause, versus the environmental collapse of the planet from climate change. In these cases it might be argued calculation should not have a defined endpoint, as to arbitrarily draw a line where outcomes shouldn’t be foreseen is avoiding a person taking responsibility for the very long-term impacts of their decisions.

There is not an easy answer to this question, and, although different types of logic, probability and economic game theory have been applied in modern ethics to the issues of utilitarian calculation, they still often clash with our moral intuitions about the extent to which people should calculate and be aware of the moral consequences of their decisions.

#### *Comparing Pleasures*

We’ve explored some of the difficulties with Mill’s understanding of higher and lower pleasures, but there are still questions as to how calculations of these qualitative distinctions should be put into practice. This issue is especially important, as, in the modern world, governments regularly make decisions as to what they think is most valuable within a given society; not just on the amount of pleasure it produces, but how important that pleasure is compared to others.

Moreover there are questions as to how qualitatively different pleasures fit in with the already discussed issues of average vs total happiness and the finishing point of calculation. For example, while competent judges might regard James Joyce as a currently higher pleasure than Adam Sandler, what if in 500 years this opinion changes, with Sandler’s films being shown in special screenings at the Louvre? Such a future seems unlikely now, but many pieces of art, sport or music are often ‘misjudged’ at the time of their creation, so should steps be put into place for calculating how pleasures might qualitatively change over time?

This question in particular may come into play when one considers the money that is often put into building concert halls, in the expectation that the value of opera or classical music will be consistent regardless of how many people attend the concerts. If one makes the argument that in the future, opera and classical music will not be regarded as higher pleasures, the costs of such projects may be too great to justify the preservation of a special place for those forms of music. Therefore, the higher/lower pleasures distinction does have real-world implications that cannot easily be ignored.

**Activity**



Look up and watch the video clip below, taken from the television show *Yes Minister* (Season 3,   
Episode 7 ‘The Middle-Class Rip Off’). Who do you think has the stronger argument? Should the arts be subsidised despite only being enjoyed by a small subsection of the population?

For each side, write down a utilitarian justification for their position. Who do you think Mill and Bentham would have sided with?

Go to **zzed.uk/9765**

🖰

**Clip**: *https://www.youtube.com/watch?v=zvNw0P5ZMbA*

#### How Might the Utilitarian Respond to Issues of Calculation?

Calculation problems are a common objection to utilitarianism, and there are at least two potential ways to respond. The first is simply for the utilitarian to ‘bite the bullet’ and agree there are calculation problems, and these will not disappear while human beings are not omniscient. However, they might still claim that human ignorance should not determine what counts as moral and what counts as immoral. Whether an act is good or bad in this way should not be held to whether human beings fail to appreciate the ramifications of their decisions. Furthermore, they might claim that even moderate attention to the outcomes of human action produces much better moral decisions than if people simply followed moral rules. Therefore, while utilitarianism is imperfect, it is still correct and the best form of normative ethics for evaluating moral action.

The other way is to argue that there are specific ways of avoiding calculation problems, and that is simply to specify properly what human beings are supposed to calculate when making a moral decision. For example, if we suppose human beings do have free will, no part of calculations could feasibly accommodate that in one’s moral decisions. I can only be understanding of the outcomes of my actions, not how my actions lead to others making free decisions that might remove the good effects of my decision. Similarly, one might look towards distinctions such as Aquinas’s division between vincible ignorance (knowledge a person should have had when making a moral decision) and invincible ignorance (knowledge they could not reasonably possess). Every moral theory necessarily, because of the limits of human intellect, needs boundaries on how it identifies moral responsibility in people’s actions. Utilitarianism is not a special case in its issues regarding moral responsibility, and so arriving at a solution is a broader ethical concern, not one against utilitarianism specifically.

Therefore, there are utilitarian responses to these issues, and many modern utilitarians have taken great care to address the shortcomings in Bentham’s and Mill’s original philosophies. A number of these will be examined in the later section on Preference Utilitarianism.

3. Does Rule Utilitarianism Collapse into Act Utilitarianism?

One common criticism of rule utilitarianism, put forward by the act utilitarian philosopher J J C Smart, is that in practice it ends being no more than a pointlessly complicated version of act utilitarianism. Consider the following rule: ‘Do not lie.’ A rule utilitarian would likely agree that following this rule is likely to produce more happiness on the whole than not following it. However, suppose there is a situation where lying would actually produce far more good than telling the truth. For example, if a seven-year-old child were to ask one of their parents on Christmas Eve whether Santa Claus really existed, it would be far better to reply ‘Of course! He’s currently winging his way over from the North Pole with a sleigh full of presents!’, than to crush their youthful naiveté with something like: ‘No he does not. There is no magic or good in the world. Life is first toil, then the grave.’

The rule utilitarian may wish to change their rules in certain cases, which would lead to rules such as ‘Don’t lie except in cases where it will preserve the innocence of youth’, or, ‘Don’t lie except in cases where it would protect national security.’ However, pretty soon the distinction between rule utilitarianism and act utilitarianism is going to become very blurred. If the morally correct *rule* which a person should follow is decided on a case-by-case basis, how is this any different to deciding the morally correct *action* on a case-by-case basis? Overall, rule utilitarianism faces difficulties when rules clash, either with each other, or with the primary utilitarian rule; the principle of utility.

**Occam’s Razor**

A philosophical principle that states the more assumptions and explanations needed to make a theory function, the less likely it is to work / be true. In other words, the simplest of any   
competing theory is preferred.



As a result, rule utilitarianism will face all the same problems as act utilitarianism. In fact, if we subscribe to **Occam’s razor**, rule utilitarianism is the *inferior* theory because it unnecessarily complicates matters. We now have both conflicts between different rules, and conflicts in calculating the utility of a particular action. On the other hand, the alternative for the rule utilitarian is to be far more inflexible. Even if following the rules in one case may have terrible consequences, it should still be followed. This was noted previously as a case of ‘superstitious rule worship’ by J J C Smart (‘Extreme and Restricted Utilitarianism’, *The Philosophical Quarterly*, Oct. 1956) and can be seen as a failure to even begin to uphold the primary aim of utilitarianism to maximise the pleasurable outcomes of any actions. In this account, therefore, neither a strong nor a weak version of rule utilitarianism is satisfactory.



There are two responses a rule utilitarian can give to these charges. Firstly, the rule utilitarian can argue that allowing too many exceptions to rules will ultimately undermine them. If people are likely to make exceptions to any given rule (such as ‘do not steal’ or ‘do not lie’) nobody will bother to obey rules in the first place. Therefore, there must, at the very least, be some limit on how much the rules can be altered. Secondly, the rule utilitarian can state that there should be some kind of safeguard rule in place which is able to prevent truly catastrophic consequences; something along the lines of ‘Break any of these rules before allowing something outright barbarous.’ The critic may still contend, however, that the imposition of some other limit or principle onto rule utilitarianism is just window dressing. There is still a fundamental issue in that rule utilitarianism cannot accurately define when rules should be broken without resorting to becoming a more complicated version of act utilitarianism.

4. Character and Motives

#### Bentham on Character and Motives

Utilitarians such as Bentham often pride themselves on their common-sense approach to ethical decision-making, but it can be argued that by excluding motive and character entirely from the picture they offer an at best incomplete, and at worst perverse, account of morality.

Certainly it is possible to find actions which produce pleasure in abundance and cause little pain but which we will still find ethically dubious.

As a consequentialist theory, act utilitarianism concerns itself only with the expected *outcome* of an ethical decision. It has no interest in *why* a person chose to do such and such a thing, or whether that person is considered good or bad by their peers. Indeed, this is a distinctive feature of utilitarianism which is not shared by any of the other ethical theories we shall study.

The problem is that by strictly focusing on consequences, utilitarianism does not capture the richness of our ethical lives. It is true that we often judge how good or bad an idea was by the results it produces, but it is also true that we can forgive someone who has caused harm but ‘meant well’ or judge a person negatively for doing the right thing for the wrong reasons.

For instance, consider the following example:

Quentin is out on a dinner date with the beautiful Astrid when he is accosted by a wizened old man collecting money for the local hospice. Quentin, who earns a considerable salary, is miserly by nature and so avoids giving up money at all costs. However, he is aware that Astrid’s late father suffered from bowel cancer and that a donation at this precise moment may well win her over. So Quentin gallantly whips out a ten pound note and stuffs it into the collection tin, ‘Here you are my good man, there’s nothing worse than cancer!’

Here we have a case where a person has carried out a good deed but for all the wrong reasons. If our instincts are to consider the motivation behind Quentin’s actions and, as a result, his character as somehow morally relevant, this suggests that Bentham’s act utilitarianism is missing something out from the ethical picture.

#### Mill on Character and Motives

Unlike Bentham, Mill does consider the importance of character, because the kind of actions we are and are not prepared to do arise out of the kind of person we are. He claims that any person who is capable of understanding what happiness is must have a good character. He also claims that doing good is one of the few things which is in and of itself good. This is because those who have the desire to do good get pleasure from doing so.

Mill is here showing, from a utilitarian perspective, the moral relevance of a person’s character and motives (because what motivates us to do things is a part of our character). If a person lacks a good character, they are less likely to increase the general well-being and more likely to increase the amount of suffering in the world. So, while in one instance a person with ulterior motives may perform a good act, that same person may at another time disregard morality entirely. It is a fairly intuitive claim; we are suspicious when someone who has been hostile and unhelpful to us in the past suddenly becomes complimentary and cheerful in our presence; our immediate thought is ‘What do they want? They are not normally so nice; they must have some hidden agenda’. We are always aware, as the saying goes, of the wolf in sheep’s clothing.

***Is all altruism really self-interested?***

We might like to think that we do good not because it brings us pleasure but because it is   
good in itself. Implicit in Mill's notion of the good character is the idea that all altruism is really   
self-interested. Aid workers do not give up their time and move thousands of miles from home to help   
build schools in South Sudan because they believe it *is* good, but rather because it makes them *feel* good. This may appear to some as cynicism, but to others as realism.



**Stretch and Challenge**

5. The Moral Status of Close Relationships

Utilitarianism asks us to take an objective stance on matters of morality. The strict utilitarian then will not assign any more value to the happiness of their frail old mother than to the happiness of a person they have never met who lives on the other side of the planet. It is the amount of pleasure for the most people which counts, rather than any notions of love or familial loyalty.

Utilitarianism demands impartiality in moral decision-making; as Mill says, ‘the happiness which forms the utilitarian standard of what is right in conduct, is not the agent’s own happiness, but that of all concerned’ (*Utilitarianism,* Ch. 2). Implicit in the principle of utility, then, is the assumption that no one person’s pleasure is more important than any other’s. What is expected is nothing less than a decision made objectively, free from all personal prejudice and bias. This means that assigning less significance to a person’s pleasure or pain because of their gender, race, hair colour, sexuality, social rank or any other aspect of their character is immediately ruled out. However, it also means that we must care no more about inflicting pain on our parents or partners than we do about inflicting it on perfect strangers. Family, friend or foe; under the steely-eyed utilitarian gaze, they are all equal.

Yet is the utilitarian right to demand this? Are close personal relationships *ethically insignificant*?

Think back to the *Trolley Problem*. Would you still switch the tracks if it were a close family member or friend who must be sacrificed to save five perfect strangers? Only an especially high-minded individual would be prepared to entirely disregard kith and kin for the sake of the utility principle. If such considerations are likely to change how a person acts then we have good grounds to argue that utilitarianism is wrong to suppose personal relationships are ethically insignificant. At the very least, the utilitarian must offer a better explanation.

#### Mill on the Moral Status of Close Relationships

Mill responds to the difficulties posed by close relationships to the utilitarian position by arguing they are not really difficulties at all.

It is a misapprehension of the utilitarian mode of thought, to conceive it as implying that people should fix their minds upon so wide a generality as the world, or society at large. The great majority of good actions are intended not for the benefit of the world, but for that of individuals… the occasions on which any person (except one in a thousand) has it in his power to [maximise happiness] on an extended scale, in other words to be a public benefactor, are but exceptional. (Utilitarianism, Ch. 2)

Mill is saying here that the vast majority of humanity rarely has the opportunity to do good ‘in general’ anyway. The ethical decisions that the average person makes will largely involve their friends, family and acquaintances; as such, most of the people affected by their moral decisions will be people with whom they have a personal relationship. It is only a few people in society – politicians and military leaders, for example – whose ethical decisions can have a widespread impact. The majority of us, Mill thinks, will never have to face a situation where we must choose between the demands of the utility principle and the demands of our hearts.

Is Mill’s argument satisfactory? Well, on the one hand, it could be argued that the average person’s actions have broader ethical consequences than Mill makes out. In the twenty-first century, a prominent blogger or an army of Twitter users can have an impact, sometimes for the better, sometimes for the worse, on the lives of people on the other side of the globe. The world was undoubtedly a much smaller place in Mill’s time, and it may be worth considering that the interconnectedness of our own age means that we have greater ethical responsibilities to one another.

One could also contend that human beings are faced with regular decisions to help those close to them or those more distant, but that simply such decisions are obscured from us by human nature to devote time primarily to relatives and friends. For example, it might be my friend’s birthday tomorrow, and I believe it is important to get them a present. However, on the way to the shops, I spot a homeless man who appears hungry and ill. At this point in my life, I do not have much disposable income, and helping feed and buy medicine for the homeless man will use up the money for my friend’s present, but provide a great deal more happiness for a suffering individual. Such situations, even in the modern world, are not uncommon, and many people are drawn towards donating their time to charity in lieu of their friends for such reasons. Furthermore, the utilitarian, as impartial, should support such efforts. Therefore, there is a tension between the utilitarian commitments to general happiness, and the way human beings prioritise making relatives and friends happy. It is difficult to discern whether human beings are simply not living up to a required standard of moral goodness, or whether utilitarianism expects too much of people.

Regardless, the main difficulty with the utilitarian position more generally is that it seems to suggest that even if we do something morally good for a close relation, we do not do so out of love or loyalty but out of respect for the utility principle. So if we provide food and shelter for our children, or if we sit by a friend’s bedside as they enter the later stages of a terminal illness, we are not doing so because we care for *them* but because we care for *the general good*. We will return to this problem again in the next section.

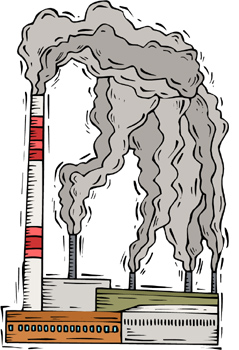
#### Discussion:

*Think about how you make moral decisions; do you calculate the happiness of each action you do? Or if concerning friends are your moral actions more instinctual and personal?*

6. Does Utilitarianism Undermine Personal Integrity?

One prominent criticism of utilitarianism (as well as Kantian ethics) was put forward by Bernard Williams. He argued that codified, or systematised, ethics such as utilitarianism require that human beings commit acts that stand in tension with their ordinary moral intuitions. What exactly is meant by this?

Well, we have examined numerous scenarios where utilitarianism has required that we act in ways we would feel uncomfortable with. Whether it be acting in strangers’ interests above our friends, or sacrificing lives for marginal gains in happiness, the very foundations of utilitarianism sometimes seem to require that we act in ways that, if left to our own moral devices, we would naturally avoid. To this extent, human beings can be said to have a ‘personal integrity’, a sense of moral self that from intuition and experience puts down certain personal principles and no-go acts. However, utilitarianism, in somewhat abstractly prioritising general happiness, overlooks such an integrity, and demands that people act in accordance with the principle of utility, regardless of what moral acts a person would not wish to perform. Therefore, utilitarianism seems to violate the ordinary moral wishes and intuitions of a person, and, in many ways, reduces them down to machines that should only compute what is pleasurable and painful in any given situation.



To take an example, consider Grace, a risk assessment officer visiting a food processing factory. Upon touring the facility she encounters a machine that is faulty, and is at great risk of breaking down in the coming weeks and severely injuring one of the factory workers. She points this out to the overseer, who tells her they were aware of the faulty machine, but it is still currently working. If it were to be replaced this would potentially lead to great delays, and would lead to many thousands of businesses and people hungry and starving in the nearby town. The overseer adds that they have ordered a new machine to be delivered, but it will take a few weeks to arrive and so it is still necessary to use the faulty machine until that date, even if it means one of the workers operating it suffers an injury.

An act utilitarian may simply think, well, if thousands of people will be affected versus one, then in this situation I should simply ignore the faulty machine. However, Grace until this point has taken her job seriously and does not want to be responsible for the injuring of a worker, even if management have agreed this is an acceptable result. The idea of simply overlooking a severe risk, even if it causes more general harm, fills her with dread, so purposely doing it seems out of the question. Many of us would also understand if Grace wished to report it. Such an act is standing by her principles, and simply looking out for the ordinary factory workers.

Yet, under act utilitarianism, such a decision is wholly wrong. It requires that Grace overlook her whole belief system, her ethics and sense of self in ignoring the faulty machinery. In this sense it undermines her purpose and integrity and commits her to act alongside principles she cannot square with her own sense of rightness.

What is more significant, however, is that this situation is not an isolated one. What Williams critically notes is that for any act conceivable, a scenario or situation can be imagined that makes that act permissible. To this extent, truly abhorrent acts such as genocide or infanticide might be justified under act utilitarianism, so long as the outcome was better than if the act was not performed. This means that any individual following act utilitarianism would be required to potentially commit these abhorrent acts, whether or not they would go against their wishes and integrity. Furthermore, it can be argued that rule utilitarianism does not escape these issues either, as it has been noted from Smart’s criticism that rule clashes mean it may collapse into a form of act utilitarianism.

***Utilitarianism, Equality and Partiality: Whose Happiness   
Should be Taken into Account?***

One extra question that can be considered in utilitarian ethics is whose happiness should be included when we calculate the impact of our actions? Should we consider animal happiness on the same level as human happiness, or should the interests of animals be not as important as ours?

In fact, Bentham was quite interested in the issue of animal rights, and campaigned for greater recognition of the suffering animals undergo in servitude to human beings. However, a modern philosopher who has developed utilitarian thought even further regarding animal welfare is Peter Singer. He argues that not taking into account the preferences and happiness of animals in ethical decision-making is a form of speciesism. In fact, the significant sentience and intelligence of many animals such as chimpanzees and dolphins means that they are deserving of a high moral status, to the point where their interests might reasonably be prioritised over those of severely disabled children, who might be more cognitively impaired. We might sum up Singer’s argument in the form below:

1. We do not eat human beings, as they possess certain characteristics such as intelligence, consciousness, ability to express preferences that makes them morally relevant and important in ethical decision-making.

2. Animals possess these morally important characteristics as well as humans.

3. Therefore, we should not eat animals.

The main issue of contention is that if Singer’s claims are accurate, then human beings should be more readily calculating the impacts on animals in their moral decision-making. Where possible, human beings should be vegetarians, inessential animal experimentation should be a thing of the past and we should no longer use animal fur in clothing, to give a few examples. The utilitarian schema for Singer should extend beyond human happiness and interest, and into all living things who might possess the intelligence and sentience to express real interests and preferences.

***‘Eating Meat and Eating People’***

Singer’s line of argument has, however, been questioned by Cora Diamond in her influential essay ‘Eating Meat and Eating People’. It is important to note that Diamond does not support eating meat – she is a vegetarian herself – but rather criticises the kind of arguments Singer gives in support of animal welfare and rights.

In particular, she asks whether it is really possible to regard sentience, or sentient intelligence, as the only way of understanding animals, and adjusting our attitudes towards them. If this were the only important factor, then it should be the case that we should have no issue with eating dead human beings, as they possess no sentience. Similarly, we generally do not eat our pets, even after death, but this is not to do with the animals’ sentience, but rather from the status they have as ‘pets’ and the moral attitudes we hold towards our pets. In this way, Diamond argues it is not possible to reduce the way human beings morally understand their relations to animals to a simple case of possessing or not possessing sentience. We hold a vast array of complex and sometimes conflicting attitudes towards animals, just as we do towards human beings, and Singer’s utilitarian justifications do not do justice to the different and important ways human beings have a moral relationship with animals.

**Further Study:** Diamond argues that the concepts of ‘pet’, ‘friend’ and ‘person’ are morally thick concepts that carry their own sets of responsibilities, requirements and duties. This potentially means that if we saw a person eating their pet dog, we would not just say they were committing a morally wrong act, we would also perhaps point out they misunderstood what a pet even was! Once you have studied virtue ethics, read through Diamond’s essay again, and see whether you agree with her assessment of these concepts and their relation to animal welfare.



**Stretch and Challenge**

Preference Utilitarianism

Preference utilitarians avoid a difficulty faced by both rule and act utilitarians.

A short story: Let me introduce myself, I am Professor Lethe and I have invented a most fabulous machine. This device, which I call Lethe’s HedoneDome, is able, through some rather complex neuroscientific whatnots that I shan’t go into, to grant its user the ability to experience a lifetime of unadulterated pleasure! Now, whether you wish to live the life of an emperor in Ancient Rome, or that of a libertine prince in an unseemly palace of carnal delights, or perhaps simply to while away your days serenely contemplating matters of the mind, I shall not judge – your choice of pleasure is yours and yours alone. What’s more, my device will erase every memory you ever had of this dreary, pain-begotten plane and let you live in the HedoneDome as if you had been born there! Imagine knowing no pain but only pleasure, and knowing neither boredom nor satiation but instead to feel that each waking moment can never be surpassed! The only catch – if it can be called that – is that you can never return. Once you enter the cave, you stay in the cave. Until the end of your days. All I need is my first volunteer, now come, step inside…

This thought experiment was first devised by Robert Nozick in his 1974 work *Anarchy, State and Utopia,* where he called devices such as the HedoneDome, ‘Experience Machines’. He intended them as a critique of both act and rule utilitarianism because both endorse **hedonism**. Nozick thought that people would not choose to abandon reality for a life of pure pleasure. He argued that people place some inherent value in *being connected to reality*. Mill cannot get around this by objecting that ‘being connected to reality’ is some kind of higher pleasure because we would think the HedoneDome *was reality* (remember: Professor Lethe said you would forget all about the real world and live as if you had been born in the HedoneDome). Any higher pleasure is just as possible in the machine as it is in reality. But if people would choose not to go into the machine this suggests they value something *other* than pleasure. That, in a stroke, defeats hedonism, because it shows that not everything can be reduced to pleasure; other things are valuable too.

**Hedonism**

From the Greek word for pleasure, *hēdonē*, this is a philosophical position which holds that pleasure is the ultimate good in life.



#### If Not Pleasure, What Should Human Beings Seek?

One of the difficulties in criticising psychological hedonism is that one has to provide alternative reasons and goals for people’s action. This is harder than it seems. Some might contend, for example, that we just desire certain objects and things, irrespective of whether they give pleasure. This is what the utilitarian Henry Sidgwick argued; that there are just objects with intrinsic qualities that we desire for the sake of them. For example, a person with a stamp collection might simply wish for more stamps to complete the collection, even if finding the missing stamps is a process filled with frustration and suffering. However, in many of our hobbies and activities, we generally engage with them at least for a sense of satisfaction, which the psychological hedonist might argue really is a form of deeper pleasure (in the case of Mill, maybe even a higher pleasure).

However, it has to be considered that the advocate for psychological hedonism may simply be ignoring direct evidence when we claim that it is not just pleasure that we seek. If, for example, I claim that I think virtue is an important aim of human beings, separate to what is pleasurable, and the psychological hedonist claims that I only seek virtue for the pleasure it gives me, then it is difficult to understand what could refute such a claim. In principle, the psychological hedonist argument may be an **unfalsifiable** argument. Why is this an issue? Well, if no amount of evidence could convince the psychological hedonist that human actions are not reducible in motive to pleasure or pain, then what sort of meaning could it possibly have? I might well argue that aliens are secretly controlling all my actions, and that anyone arguing against this view is simply repeating what the aliens have told them to say.

What this means is that the psychological hedonist cannot simply just argue everything is reduced to pleasure or pain, whether we believe that or not. Evidence needs to be given, whether biological or philosophical, or there are reasonable grounds to suggest that there may be a variety of ends to human conduct, not just pleasure and pain.

#### Discussion:

*Would you step inside Professor Lethe’s Pleasure Cave? Why might this thought experiment be   
problematic for utilitarians?*

#### From Pleasure to Preference

Later utilitarians were aware of problems such as this and decided that rather than talking about the maximisation of pleasure they should talk about the maximisation of *preferences*.

A preference is anything we have an interest in pursuing. Sometimes, it might be the case that satisfying a preference does not bring somebody pleasure. For instance, satisfying our preference to stay in reality might not always bring us the most pleasure but it is our preference nonetheless. Perhaps this is why some people prefer sobriety; they know getting drunk might be fun, but ultimately, they think it isn’t *real*.

An important point to raise here is that preference utilitarians do not necessarily think it is just human beings who have preferences. All sentient beings (those capable of feeling) have preferences too, even if it is just the preference not to be killed.

**Preference Utilitarianism**

Holds that the right action is the one which satisfies the greatest number   
of preferences.



So, Nozick’s objection can be overcome and, as you shall see when you come to study Applied Ethics, **preference utilitarianism** offers other significant advantages over traditional act and rule utilitarianism as well as some startling conclusions about how we should live.

#### Advantages of Preference Utilitarianism

The main advantage of preference utilitarianism is that it avoids many of the issues observed with viewing everything in terms of pleasure or pain. Rather than having to analyse all human actions on this reductive level, one is simply able to admit that human beings have a myriad of motivations for actions in their lives, and these should be accepted as being important regardless of what relationship they may have to pleasure or pain. This helps cover significant areas of ethics where actions may be considered right without them bringing an excess of pleasure in any case. For example, spending more time with my terminally ill pet may bring me more pleasure as I enjoy their company for longer, but it would bring the pet significant pain. If one considers all the preferences in that situation, including both mine and the pet’s wish to not suffer in great pain, one could more easily arrive at the conclusion that I should euthanise my pet, no matter how much pain it might bring me.

In particular, it solves problems such as Nozick’s ‘Experience Machines’, for, while being in one could bring me much more pleasure than if I lived in reality, it can be argued that despite this I have a preference for ‘real’ experiences (alternatively, I could have a preference for artificial reality, this is not excluded). Another significant advantage this may bring is that the charges against Bentham for encouraging social manipulation are avoided. Legislation does not have to be brought in to nudge people in favour of maximising general happiness. Rather, one can simply consult what the preferences of each person would be, regardless of their relation to pleasure or pain, and so a society can be formed that works based on people’s real wishes, instead of what they believe would bring them the most pleasure. This may still be contentious, however, as will be examined later, for one can question whether people’s preferences are really rational, or whether one would still have to manipulatively encourage more positive preferences in any given society.

Perhaps the greatest advantage of preference utilitarianism, however, is that it avoids a number of the problems of calculation. For the difficulties associated with accurately comparing and quantifying pleasure disappear; all we have to do is focus on the different preferences people may have. There are naturally still difficulties with this, as it may be hard to understand everyone’s preferences at a particular point in time, but it allows for the effective narrowing of the scope of utilitarian measurement. For example, if in a particular society I was required to legislate on the different kinds of rides in a theme park, I could simply offer a referendum-type vote to all people to discover what kinds of ride they would prefer, rather than attempting to judge what kinds of ride would bring them the most pleasure.

It is important to understand why many utilitarians prefer a different approach to the principle of utility. While initially seeming simple, there are a great many counterintuitive implications of it, and a more nuanced approach, such as maximising preferences, may be easier to understand, implement and align closer with people’s ordinary moral intuitions.

Issues with Preference Utilitarianism

#### Should All Preferences be Satisfied?

One criticism that can be levelled at preference utilitarianism is that it could make it morally right to allow people to satisfy some downright dangerous desires. If my sole aim in life is to stick pins into every square centimetre of my body, then, according to a preference utilitarian, it would be morally right to allow me to satisfy that preference. Yet surely such a preference is evidence of some underlying mental illness, or, at any rate, is hard to consider as a good way to live. The case is even stronger for those with severe depression whose preference is to commit suicide. Examples like this raise the question of whether it is good to satisfy *all* preferences or only some. If the latter, the preference utilitarian faces the difficult, if not impossible, task of deciding which preferences we can call good. [*Stanford Encyclopaedia of Philosophy*, Consequentialism, Sinnott-Armstrong, W]

A more everyday example might be if a 16-year-old teenager wishes to eat ice cream every day, despite their parents warning that it will rot their teeth. At this point, whose preference should be allowed? The teenager is old enough to know the risks of eating ice cream every day but perhaps cannot comprehend what they mean, while the parents’ preferences are that the freedoms of the teenager should be restricted. At some point it can be argued there has to be a principle established that details what principles might be regarded as reasonable. For if a preference is built out of a delusion, or ignorance, should it really be fulfilled? This leads to the greater question, of how we can possibly altogether balance people’s different preferences.

#### How Does One Judge Which Preferences Should be Prioritised?

Examine the following example:

Terry is greatly opposed to all forms of immigration and emigration. In fact he believes this to such an extent that he wishes that everyone in the world was deported to their country of origin, thereby eliminating any possibility of racial tensions and xenophobia. He justifies this by arguing that if everyone’s preferences were satisfied about where they would want to live, then everyone would live in the richer countries, destabilising them and ruining the prosperity particular countries have worked to achieve over the years. Furthermore, he argues everyone has a preference really to remain in their own culture, and so would not really express a real preference to live outside where they were born, if there was no choice in the matter.

How should we approach Terry’s claims? On the one hand, it seems to be an extreme position of racism and xenophobia, but the preference utilitarian may claim that if there are enough Terrys we should satisfy their preferences, even if a minority oppose them. Therefore, at first it is apparent that preference utilitarianism, perhaps more than hedonistic utilitarianism, has difficulty dealing with issues of rights and protection of minorities against opposing preferences. Furthermore, there is a more nuanced difficulty here: should Terry’s preferences be satisfied even if there is no causal connection between him and the people it affects? Many people have opinions about how the world should be, and it is questionable whether one’s preferences should extend beyond one’s own interests.

For example, I might wish for peace between conflicting countries in the Middle East, and many would argue such a goal is a positive one. To this end, one might argue that my preference should be satisfied. But what if, in the warring countries in the Middle East, there is a majority of people who would prefer to continue fighting? Or what if there are interventionist countries around the world making lots of money selling arms to these countries, who wish the war to continue? What seems to be needed is some extra set of principles that determines what preferences are acceptable, thereby eliminating the central aim of preference utilitarianism in trying to analyse acts in terms of their consequences.

#### The Quantification of Preferences

We can ask at this point, has preference utilitarianism really presented itself as a viable alternative to hedonistic forms of utilitarianism? In fact, it can be argued that, by removing a strict impartial understanding of pleasure and pain, it becomes much more complicated, and critics have argued similar complications exist when we try to quantify preferences.

For example, if I was strongly campaigning for a new library to be built in my local residential area in the belief it would help improve the welfare of local children, should my preference here count above the individual weakly opposing it because they find children a bit irritating? It is difficult to judge to what extent the strength of a particular preference should count in its favour. Yet if it does not count, one might find that one’s moral decisions and beliefs are continually hampered by a miserly section of the population, who simply wish for no general preferences to be fulfilled.

Similarly, what if there is an enthusiastic supporter of local development works who has thousands of preferences about where new buildings should go, and the people they might benefit, versus the lackadaisical, ambivalent individual who has no preferences but the location of his own home. Should we fulfil all the preferences of the enthusiastic individual, regardless of whether they overwhelmingly outnumber everyone else’s? These questions of quantifying preferences, as we have noted, exist in similar forms of hedonistic utilitarianism, but they are still pertinent criticisms, and perhaps even greater once one removes a more universally understandable measure of pleasure and pain from the theory.

***Alternative Forms of Utilitarianism***

Through this section, forms of utilitarianism have been examined that focus on natural properties as the goods to be maximised. However, other forms of utilitarianism exist which focus on maximising non-natural properties. What does this mean?

The philosopher G E Moore, from his criticism of naturalistic ethics, argued that what is good cannot be reduced to something similar to pleasure. Rather, it is just an ideal, grasped through human intuition, that if maximised could bring about generally good states of affairs. In this way, Moore argued that utilitarianism should focus on maximising these non-natural intrinsic goods, which not only included what intuitively could be considered to be moral, but also things such as beauty, and loving relationships. This is sometimes called ‘ideal utilitarianism’.

However, one potential criticism of such a view comes from J J C Smart, who gives an example of the ‘deluded sadist’. Imagine that there is a world where only one person exists. However, this one person is severely mentally ill, and is deluded into thinking that all around him in this world are people being tortured, and he is not capable of preventing this torture from occurring. Smart asks which of the two situations below would be better overall:

1. The deluded individual is continually horrified at the images of others getting tortured, and suffers a great deal throughout his existence.

2. The deluded individual actually experiences a great deal of pleasure when watching others get tortured, and so for all intents and purposes is quite happy with his existence.

We can now ask, which situation is morally preferable? Instinctively one might be inclined to argue for 1, as the thought of people gaining pleasure from the suffering of others is something disconcerting in the moral framework of our world. Such a belief also would match Moore’s understanding of good. Here we cannot simply equate what is good with what is pleasurable; we understand that even if there is not pleasure involved, a situation can sometimes still be good.

However, Smart asks the reader to keep an open mind. He notes that within our lives we are uncomfortable with situation 2 as we normally associate those who take pleasure in others suffering with real sadists, who do inflict harm on others. But in this situation, the pleasure of the deluded individual causes no harm, and we should really see situation 2 as the morally better state, for why would one want the deluded individual to simply suffer in silence? As Smart states:

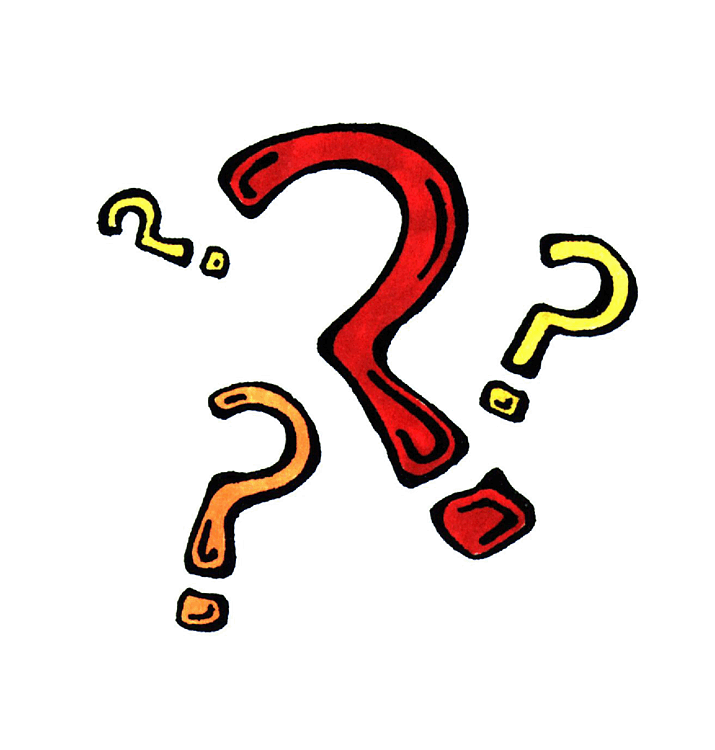
*If we lived in a universe in which by some extraordinary laws of psychology a sadist was always confounded by his own knavish tricks and invariably did a great deal of good, then we should feel better disposed towards the sadistic mentality.* (J J C Smart, *Utilitarianism: For and Against* pp. 25–26)

However, many disagree with Smart here, arguing that we do not simply measure what is good and bad according to what is pleasurable, and that in fact his example is evidence that intuitively some things are just bad! How you view the deluded sadist’s situation, therefore, may be evidence for or against hedonistic utilitarianism.

**Note:** The question of whether good can be reduced down to natural properties is covered further in the Meta-ethics section.



**Stretch and Challenge**



**Utilitarianism: Quick Quiz**

1. How does Bentham define utility?

2. Identify three factors that must be taken into account when calculating pleasure with Bentham's utility calculus.

3. How is a higher pleasure to be distinguished from a lower pleasure?

4. What is meant by consequentialism?

5. Define hedonism.

Ethical Theories 2: Kantian Deontological Ethics

Introduction to Kant

Immanuel Kant (1724–1804) is one of the most important philosophers in the Western philosophical tradition, and has proved greatly influential in nearly all major areas of philosophy, including epistemology, metaphysics, ethics and aesthetics.

**Anthology Text:**

Kant, I (1785) *Groundwork of the Metaphysics of Morals*



The major work that will be focused on in this section is the *Groundwork of the Metaphysics of Morals*. Written in 1785, it was Kant’s first major work on ethics, and highly ambitious in its aims. In fact he aimed to not only put forward his view on major ethical issues of the time, but provide an account of fundamentally what morality was and why it applies to all human beings!

This was, and still is, a tall order, and, despite Kant’s work proving influential, it still contains many flaws. Nevertheless, there are many deontological philosophies still being put forward today, sometimes built upon Kant’s work. People who draw inspiration from Kant today are often called Kantian ethicists.

#### *The Projects of Rationalism and Empiricism*

Kant wrote during a period of history now called the Enlightenment; a time when many philosophers began questioning the theological dogma of previous centuries, and putting forward reason as the principle by which philosophy and science should be conducted. Kant was no exception to this, as will be seen, and he attempted to derive much of his philosophy using reason alone, including ethics.

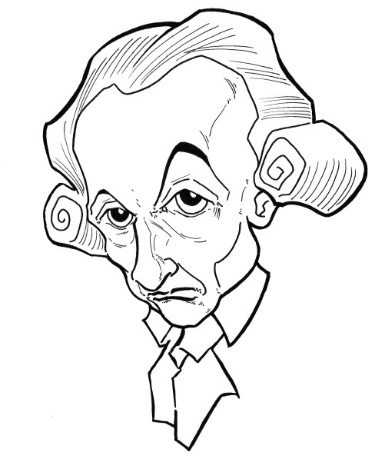
However, there was a significant debate during Kant’s lifetime about how human beings come to truly know things about the external world (one that continues to this day). On the one hand, there were the rationalists, those who believed real truth comes from reasoning on the world, independent of sense experience. To this end, truth was innate; one could discover it whether or not one had relevant experience that could confirm it, and fields such as maths and geometry were used as examples of truths that were independent of sense experience. On the other hand, there were empiricists, who believed that human beings were born a ‘tabula rasa’ (blank slate) and that all knowledge and truth comes originally from sense experience of the world, arguing that if one followed the psychological connections between ideas closely, one could identify the experiences they arose from.

Kant attempted to find a path between these two competing philosophical beliefs. He did not believe in the pure rationalist ideal and recognised that experience had a significant element in forming human knowledge, with the independent world (the noumena) out of reach of perception. However, he equally believed that there were structures of thought that governed how the world was experienced, and could not be understood simply from the content of experiences themselves. These structures were governed by certain rules, and Kant believed that philosophy had the task of discovering these rules and how they impacted the way human beings come to know things, including moral truths.

#### The Centrality of Reason

At the centre of Kant’s philosophy was reason. He believed that ethics could not be developed through an appeal to emotion, pure experience or natural instinct, as these were unreliable. Each person has innumerable different motives and desires for acting, and so if we appeal to any of these sources Kant believed we’re bound to end with at best conflict, and at worst a morality that is little more than opinion. Kant, therefore, believed that a true account of morality had to be universal; it had to apply to anyone regardless of their own desires, experiences or aims, and so he looked to reason as a faculty that could discover the moral rules that guide humanity. The process in which we work through what human beings should do is, therefore, an abstract one. We examine the form and nature of morality, and from that reason what normative ethical laws might be deduced.

Kant, therefore, does not look to experience, as Mill does, to justify his views, but rather, as part of a deeper philosophical project based on reason, aims to derive morality from basic first principles. This means that Kant, perhaps more than any other modern philosopher, presents a fully systematised account of normative ethics, parts of which at first glance can appear quite technical and complicated. Working through this course companion, it is important to understand each step Kant makes in justifying his presentation of what he views as the central defining principle of normative ethics: the **categorical imperative**. This imperative, he argues, allows human beings to work out what rules and obligations they have a duty to follow, such that correct reasoning in any situation can determine what one **ought** or **ought not** to do.



‘Two things fill the mind with ever new and increasing admiration and awe, the more often and steadily we reflect upon them: the starry heavens above me and the moral law within me.’

Yet despite Kant’s systematic approach to the categorical imperative, many have   
criticised different elements of his overall moral philosophy. Some contend generally   
that it is not possible to determine what is moral simply from reason alone, while some argue   
more precisely that the categorical imperative itself results in unreasonable or unintuitive moral   
laws. Therefore, whenever Kant makes a philosophical claim, it is important to assess it fully and judge for yourself whether it can be justified within his overall system of ethics, and whether it holds up against the common moral intuitions human beings possess.

The first place to start is Kant’s interpretation of the moral good itself, and what it means for someone to act in a morally good way. This forms the basis for his understanding of morality as universal law, and leads to his argument for the categorical imperative as the fundamental principle that guides us in working out these universal laws.

The Good Will

#### What is the Good Will?

The first question we can ask when trying to understand ethics, is ‘What makes an action good?’ Many different answers have been given to this question, and, for any moral philosopher, beginning to answer it opens many more questions than they would likely prefer.

For example, say I heard on the radio that a member of my favourite boy band has become seriously ill and needs a kidney transplant. Unfortunately he has a particularly rare blood type which means there are very few potential donors and so, from the doctor’s prognosis, it seems unlikely that the celebrity will survive longer than a few months. By chance, however, I discover that I happen to have the same blood type, and since I cannot bear to think of a world where my favourite boy band cannot keep making music, I decide to donate my kidney to the suffering individual, receiving much praise from the press in the process and ensuring I can enjoy the boy band’s music for years to come.

Was my action good? Mill would argue yes. Whatever my intentions, my actions had good consequences, bringing greater happiness to all involved, and so there is no question as to whether I should have or should not have donated my kidney.

Yet Kant would potentially argue differently. For, we can ask, what was my motivation in donating my kidney? In the example, it was so I could keep hearing the boy band’s music, not because I considered that it was a general duty for me to donate my kidney to those who are in need. If the potential recipient of my action was not my favourite boy band member, but a random member of the public, I most likely would not have offered my kidney. This raises an interesting question: does the ‘end’ of my action influence whether or not my action should be considered good?

Kant strongly argues yes. When we consider the numerous different ends of our actions, he contends that they cannot be considered good ‘without qualification’. When we look at any end of action such as happiness or virtue, it is possible to argue that in certain situations acting towards these ends can result in a bad action. For example, if I donated my kidney to a fascist dictator who regularly tortured and oppressed those under his rule, as I sadistically found happiness watching others suffer, we would be less inclined to consider my action good. This means, considering the original example, that we cannot view the donation of my kidney as a wholly good act as it was done for my own happiness, in ensuring my favourite boy band kept making music. In this case it was good, but the pursuit of happiness may well also lead to morally bad acts.

So what for Kant is the right end or reason for moral action, if not a particular identifiable end? Well, Kant argues the only measure of whether an action is good is if is performed from a **good will**. The good will is the source of all good, is good without qualification and ensures that when we are acting, we are doing so for the sake of doing good, not for some other reason or motive. In this way we can say the good will is **intrinsically good**.

**Deontological**

Comes from the Greek *deon*: that which is necessary (i.e. an obligation or duty). Combined with -ology we get the science of that which is necessary or the science of duty (just as psych*ology*is the science of psyche,   
the mind).



We might initially respond that this seems a bit circular. Of course acting in a morally good way involves wanting to will good things, but Kant offers a more precise definition of the good will which influences the rest of his moral philosophy; a good will is one which acts of duty. In the case of the original example, this separates the person who donates a kidney because it’s the right thing to do, from the person who donates a kidney because they wish to hear more boy band music.

#### Discussion:

*Compare the good will to psychological hedonism. Do you think there is a part of you which does good   
things because you simply believe you should do them? Or are there always other motives at work?*

#### Good Will, Moral Worth and Duty

We can contrast Kant’s idea of the good will more concretely with utilitarian ethics previously studied. You might remember in the case of Mill that although good motives might help to produce good overall actions, they had no indication on the ‘moral worth’ of an action. Rather they only have value based on the good consequences they produce.

However, for Kant, acts have moral worth insofar as they are performed out of a sense of duty. Their consequences are not relevant, and so Kant’s ethics are best described as **deontological**. This involves following what one has a duty to do, or more simply what certain moral rules dictate.

More precisely, Kant himself defines duty as ‘the necessity of acting out of reverence for the law’. When we consider what duty is for Kant, we are considering what moral laws and obligations govern our lives. This means that if we act out of duty, we are acting out of an acknowledgement of the importance of moral laws and obligations in our lives. We could not claim that we had respect for our duties if we recognised what they were, but chose to act in self-interest instead.

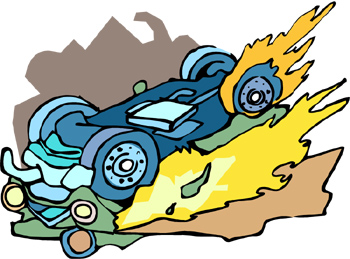
Overall, therefore, Kant’s idea of what is good involves a distinction between a pure moral sense of an action, against the other various motives and reasons we might have for performing it. While this might seem alien at times, it can be understood intuitively when looking at other forms of moral action. For example, a fighter pilot in World War II may well say he put himself in danger out of a duty to his country, or a lifeguard saving a drowning person might argue they did so also out of moral duty to save lives. When we contrast these examples with the case of donating my kidney, we can consider that there is a distinct difference between acting out of what one considers to be good, and acting because I wish some good to come about for myself. For Kant, only the former can hope to be useful when trying to understand what might be universally morally good. Not everyone will share my love for boy band music, and nor ought they to. It is in this important sense my own desires and motives cannot hope to be the basis for a moral rule that applies to everyone impartially.

Therefore, when examining Kant, keep in mind that in his discussion of the good will he is not looking at examples of what certain people might consider to be good, but rather what could be considered to be unconditionally good; that which does not rely on us holding some common desire or end within the world.

Duty, Reason and Self-interest

#### Kant’s View of Human Beings

The discussion of good will gives a deeper insight into Kant’s overall view of human beings. On the one hand, we have seen that human beings all have individual desires and motives, based on what we each consider important as ends of our actions. Yet there is also a sense in which human beings have an independent idea of what might be the right action in any particular situation, which we might follow despite it conflicting with our desires and motives. For Kant this idea of the ought or duty is revealed by reason, but is often subverted or influenced by our non-moral desires. Human beings are, therefore, imperfectly rational; while they have the ability to work out what the right moral action is, this is often tempered by their own self-interest, the root of their non-moral desires. Were human beings conversely perfectly rational, they would always act according to their duty, while if they were purely self-interested they would never choose performing the right moral action over their own desires.



Therefore, in any ethical dilemma there is a tension between the reasonable part of ourselves, and the part motivated by self-interest. If I see a person trapped in a burning car, my reasonable self would speak up and say I had a duty to save the person whose life is threatened, while the part of me motivated by self-interest would be hesitant, fearing the pain and threats to my own life from the rapidly advancing flames.

#### Duty, Rationality and Autonomy

We can highlight a few interesting aspects of Kant’s philosophy revealed by this tension between rationality and self-interest. One, that Kant notes himself, is that any completely rational being could never be motivated by duty, for they would not have any self-interest guiding their actions, and so could never choose to be rational over being self-interested. It simply would not be a consideration that they could choose some other end to their actions than what was rational and good. Similarly, a being who is completely self-interested would never realise the ability to act out of duty; there would not be a rationality that could allow them to choose any course of action over their own desires.

These extremes reveal the unique position of humans within Kant’s thought. He stipulates that it is a **postulate** (something necessary as a basis for philosophical discussion) for ethics that human beings have **free will**, or **autonomy**, and that the ability for human beings to reason and act against their self-interest when they identify it to be in line with their duty demonstrates that human beings do have an important degree of autonomy not enjoyed by animals. However, this does not mean of course that human beings are necessarily always free to choose their actions, and similarly it does not mean that a person who acts according to what is their duty is always freely choosing that action.

#### Acting in Accordance with Duty vs Acting Out of Duty

Consider again the example explored earlier of the kidney transplant. We explored how Kant would argue that I had not been acting out of good will when donating my kidney to my favourite boy band member, such that my action would not have moral worth despite it perhaps still being the right action in the situation. This highlights the difference between simply acting in accordance with duty and acting out of duty. The former is where a person does what would be the right action, but out of the wrong kind of motive. The latter is where a person acts according to what they identify is their duty, and not for reasons of self-interest or personal desire.

What is most significant, perhaps, is that Kant radically proposes that only actions out of duty have moral worth. This means effectively that people who arrive at the right action through the use of impartial reason should be morally praised more than those who act for personal reasons, including empathy, compassion or concern. Rather than wish for more music, or the recovery of my favourite boy band member, I should have dispassionately thought about what my duty was (in this case perhaps still donating my kidney) and act on whatever I determined to be the right action.

Many might argue this is a strange conclusion to draw. Would the person donating his time in an Alzheimer’s home be any less praiseworthy because he does it out of compassion rather than a sense of duty? However, we can here state that people’s sense of duty and their interests can coincide, and it is no fault if they do. What Kant is saying is that in all moral action, people should have an awareness of what their duty is, as well as what their interests are. For there may be (and almost certainly will be) times when people’s interests and duties will differ, and so long as a person only acts according to their interests, they may miss important times when duty dictates an alternative course of action. These situations will be explored in great detail when we turn to considering the categorical imperative.

1: Categorical and Hypothetical Imperatives

#### The Nature of Universal Laws

We’ve seen how Kant defines the good as acting out of one’s duty, which means using one’s reason to determine and act upon the correct moral law. Yet we can still ask what form this moral law takes.

If in any situation, a person should be able to use their reason to work out what the moral law is, separate from their own interests and desires, then a few things can be said about the kind of moral laws Kant is looking to derive. In particular, they require two properties:

1. They are **objective**. This means that they are not constrained or influenced by a person’s own views and feelings. If moral laws were dependent on holding some desire or belief then they could hardly be available to all people through the use of reason.

2. They are **universal**. They apply to everybody equally, and are applicable to all ethical dilemmas. If they did not it would not be possible to know in a particular situation whether or not a certain moral law would apply or not.

So when we consider what might be a correct moral rule, Kant contends that it has to independently and consistently apply to all individuals, irrespective of the circumstances they find themselves in. In other words, any moral rule has to be able to be rationally ‘universalised’, meaning that there are no contradictions in imagining a world in which everyone follows that particular moral rule.

Therefore, in considering the properties that a legitimate moral rule requires, Kant is able to derive a fundamental principle governing how we can derive moral rules in any situation. However, he also arrives at the same conclusion through analysing the idea of a **moral imperative**; a command describing what one ought or ought not to do.

#### Hypothetical and Categorical Imperatives

Kant draws a distinction between categorical and hypothetical imperatives:

All imperatives command either hypothetically or categorically. The former expressed the practical necessity of some possible action as a means to achieving something else that one does or might want. An imperative would be categorical if it represented an action as being objectively necessary in itself without regard to any other end (Groundwork, Ch. 2)

**Hypothetical imperatives** usually look like this: ‘You ought to do x if you want to get y.’

They command us to do something *if* we want something else. They are *conditional* commands that depend on us having a certain aim in our actions. For example, we might say, ‘*if you want to be healthy, you should eat an apple every day*’, with the understanding that eating apples is only good if we aspire to be healthy. There may be those who for some strange reason wish to be ill, and so do not believe eating apples will be good for them. Hypothetical imperatives, therefore, express some form of self-interested reasoning, not the kind of objective reasoning which Kant requires to justify moral laws.

**Categorical imperatives** are usually of the form ‘You ought to do x.’

They command us to *just do it*. This defines them as *unconditional* commands that do not depend on us having a particular desire or aim to our actions. For example, we might say ‘you ought not to lie’ and this is meant to mean irrespective of the consequences produced by telling the truth. Therefore, we can identify in categorical imperatives both objectivity and universality, for they do not refer to any particular desire or aim and apply regardless of situation or context. For Kant, therefore, categorical imperatives are the only suitable way of working out our universal moral duties, and the only guide to actions that will truly possess moral worth.

**Activity**

Identify whether each of the statements below could be considered a hypothetical imperative or a   
categorical imperative.

i) Put your hands above your head and step away from the vehicle!

ii) Buy one 12’ pizza at Plato’s Pizzeria and get the second absolutely free!

iii) If you want that leather-bound set of Kant’s collected works for Christmas, you’d better finish your Philosophy homework.



iv) Everyone should own an Apple product.

v) You oughtn’t wear white trousers if you’re going for a curry this evening.

vi) Listen, you really need to shut up.

vii) If you love me then you will give me a child.

vii) Worship me!

2: First Formulation of the Categorical Imperative

#### What Categorical Imperatives?

Kant, as we have seen, concludes in different ways that the only way of reliably calculating our moral duties is through categorical imperatives, and that the main guide to whether a duty is correct is whether it can be rationally ‘universalised’. One should only behave in a way where one can envision a world where everyone consistently would rationally behave in the same way.

In fact, Kant believed there was only one categorical imperative, although he put it forward in a number of different forms. The first form goes as follows:

‘Act only in accordance with that maxim through which you can at the same time will that it become a   
universal law.’

This is sometimes known as ‘The Formula of the Universal Law of Nature’; it refers to how universal laws (i.e. laws which apply to everybody) of morality must be kept in mind when making moral decisions.

#### The Universalisability Test

Kant’s first formulation of the categorical imperative gives rise to a particular procedure for making moral decisions which is sometimes known as the universalisability test.

The core idea is that when we make any moral decision we must ask ourselves whether everybody else could do the same thing in similar circumstances. In Kant’s language, we are making a maxim and working out whether it could become a universal law.

For some actions, this is clearly impossible. If I like the look of your new mobile phone and decide that I would quite like to steal it, Kant would immediately ask that I stop and consider the following question: ‘Could you honestly recommend the maxim that in any circumstance where somebody feels like stealing, they should go ahead and do it?’

**Maxim**

Another word for a general rule or principle, e.g. ‘Don’t drive over 30 mph in the town centre’, ‘Don’t run   
in the corridor’).



Now, I might well turn round and say ‘Whatever Kant, I don’t care, I want that phone’ but this simply would not work under the categorical imperative. For if everyone were permitted to steal when they felt like it, the entire notion of private property would soon disappear. In fact, people would probably not bother owning anything if it was only going to be stolen, so pretty soon there would be nothing to steal. Some might even contend that a world in which everybody just stole everything they wanted would prevent anyone from reliably trusting and engaging with other people, stopping the self-preservation of the human race. In other words, that maxim ‘It is permissible to steal whenever you feel like it’ is self-defeating. We can’t conceivably imagine a world in which such a moral principle should be acted on by every moral agent.

On the other hand, a maxim such as ‘It is not permissible to torture small children’ can easily be universalised. If nobody ever tortures small children, no problems are likely to arise. Certainly there is nothing contradictory or self-defeating about that rule, and in fact such a rule might be in everyone’s rational interest as it would guarantee as children that they wouldn’t be tortured, or that their own children would not be tortured by others. The opposite, however (‘It is permissible to torture small children’), could quickly defeat itself, for perhaps people would stop having children altogether if they knew they were going to be tortured. There is also another reason that Kant would have a big problem with torture, as we shall shortly see.

#### The Golden Rule

Some have compared the categorical imperative to the Golden Rule, a traditional moral idea of ‘only do to others what you would have them to do you’. However, there are a number of key differences, the main one being that the Golden Rule arises out of self-interest, whereas Kant envisions the categorical imperative as arising out of pure rational consideration. Therefore, under the Golden Rule, a person who is a masochist, and enjoys pain, might argue that it is acceptable to throw rocks at each other, but under the categorical imperative, it is unlikely a moral rule could be formalised from this proposition, as not everyone enjoys having rocks thrown at them, and universal moral rules require that we do not give preference to anyone’s individual desires.

**Activity**



Identify whether the following activities could be universalised without contradiction:

a) Stealing bread to feed starving children

b) Queuing for the cash machine

c) Maintaining a heroin addiction

d) Having a roast lunch every Sunday

e) Vandalising urinals

f) Supporting Manchester United

g) Working as a doctor

h) Lying about your age and appearance on an online dating site

#### Perfect and Imperfect Duties

In the *Groundwork*, Kant divides duties into perfect and imperfect categories. Perfect duties are those such as ‘Do not kill’, ‘Do not steal’ – their universalisation involves an obvious logical contradiction. This is often called a contradiction in conception (or conceivability), for we cannot imagine a world in which as a moral rule the core concepts retain any sense of meaning or coherency. Kant’s example of a perfect duty is ‘Do not make deceitful promises’. In particular he states:

Should I be able to say to myself, ‘Every one may make a deceitful promise when he finds himself in a difficulty from which he cannot otherwise extricate himself?’ Then I presently become aware that while I can will the lie, I can no means will that lying should be a universal law…

What does Kant mean here? Well, while we might be able to make a false promise based on our own self-interest at a given point in time, the same thing cannot be applied to everyone, for if everyone made false promises, there would be no way of trusting that anyone would keep their promises, making the entire concept of promises meaningless. Therefore, a duty to only make promises one can keep is binding at all times and so an example of a perfect duty under Kant’s ethics.

On the other hand, imperfect duties relate to those maxims which do not involve any logical contradiction but are nonetheless undesirable for a rational being. This is often called a contradiction in will, rather than a contradiction in conception. For instance, there is no logical contradiction in universalising the maxim ‘When in need, feel free to urinate on tombstones’; society would not collapse if individuals popped over to the cemetery whenever they needed the bathroom. It is hard to imagine anyone wanting to live in such a world where everyone is happy to defile the graves of each other’s ancestors, however. Thus Kant would argue we have an imperfect duty to ‘Act respectfully towards the dead’.

#### Imperfect Duties and Moral Responsibility

Kant distinguishes between duties to ourselves and duties to others, as, while intuitively we might view moral obligations as wholly to other people, if a person does not aim to improve themselves then in many ways they cannot fulfil the various duties they might have towards others. For example, we might say that the doctor has a duty to research diseases beyond their required medical reading so that they can be aware of new methods of helping patients and ensure the growth in medical knowledge has not gone to waste.

In fact, Kant gives an example of a particularly interesting imperfect duty to ourselves that covers such examples. He argues that human beings have an imperfect duty to develop their talents, and not let them go to waste through laziness. This is important as the growth of knowledge and ability in people is essential for the core functioning of human society. If we did not have builders, houses could not be built; therefore, those with a talent for building should develop it accordingly. However, if we try to universalise such a duty, it becomes difficult to justify. For the builder surely should not spend all their time getting better at building? This would prevent them from even building things in the first place, and mean they could not fulfil other duties they might have. So the duty to develop one’s talents is an imperfect one. An individual is obliged throughout their life to avoid laziness in their pursuits, but not only dedicate their lives to this obligation.

#### Discussion – Examples of Perfect/Imperfect Duties

*We have seen how Kant gives the examples of not making false promises and developing one’s talents as examples of perfect and imperfect duties. However, he also gives one more example of each kind of duty listed in the table below. By yourself or in groups, analyse each new duty and discuss why you believe it is a perfect or imperfect duty, writing down your answers.*

|  |  |  |
| --- | --- | --- |
|  | **Imperfect Duty** | **Perfect Duty** |
| **To ourselves** | Cultivating one’s talents | Not to commit suicide |
| **To others** | Charitably helping others | Not to make false promises |

3: Second Formulation of the Categorical Imperative

Kant does not stop at the first formulation of the categorical imperative, and argues that it is possible to derive a second form that illuminates further the kinds of duties and obligations people have to other people. In fact, the second form might be the most influential aspect of Kant’s moral theory, as it confirms many ordinary intuitions human beings have about treating other people, and helps in discerning what forms of moral action might be acceptable insofar as they affect other people’s lives. He presents the second formulation as such:

So act as to treat humanity, both in your own person, and in the person of every other, always at the same time as an end, never simply as a means.

This is sometimes known as the ‘the humanity formula’; it refers to how humanity should be treated when making moral decisions.

One of the problems we encountered with act utilitarianism was its dubious stance on human rights. As a relativist theory, all options remain on the table, even if the very thought may make us squirm. So, to take the classic example, torturing the innocent to save the multitude may be permissible if it promotes the greatest happiness for the greatest number.

Kant takes the opposite view. Certain actions are *never* permissible. It is this aspect of his philosophy, the so-called ‘humanity formula’, which has had the most enduring influence on Western thought.

The crucial idea here is that people can never be used as a means to an end. Torturing an innocent man to stop a twisted sadist from setting off a chemical weapon in Piccadilly Circus can never be justified for Kant, because somebody is being *used* as a tool for some other purpose.

There is an intuitive appeal to this idea; we don’t like the idea of simply being used for someone else’s gain, neither do many people take comfort in the idea of treating someone as a mere object, similar to any other. For example, if someone offered you money to wash their windows, then once the job was done refused to pay, and said they simply wanted to get their windows clean, you would probably be outraged, no matter whether they made a convincing argument that everyone was happier by withholding payment. Everyone has a strong belief that they should be respected, and the second formulation reinforces that belief.

Kant recognised this, and thought he could explain *why* people feel this way. People feel like this, Kant said, because they are rational beings, and rational beings deserve dignity, they deserve respect, and that means *always treating them as an end in themselves*.

#### Means and ‘Mere’ Means

We can ask for more clarification, though, from Kant about what it means to treat someone as a means. For example, if I order a taxi to work in the knowledge I am simply using a taxi driver to get from one place to another, does this count as treating them as a means? Similarly, is my boss right now treating me as a mere means towards making a profit for himself? It can be noted that what counts as a means is ambiguous, and so Kant takes care to differentiate between treating someone as a means, and treating them as a ‘mere’ means. The former is acceptable so long as interactions between the people involved are accepted as demonstrating mutual respect and dignity for each other in their actions, whereas the latter involves at least one person fundamentally denying this respect and dignity for the others.

Therefore, in our example, if I took the taxi, but refused to pay the driver, I could be said to be using the taxi driver as a mere means. However, if we both agreed a price for the journey, and I paid what was owed to the driver at the end of my journey, I could reasonably say that I had treated the driver as an end and so had not violated the second formulation of the categorical imperative.

Nevertheless, there are still questions that can be raised about where the line between means and mere means is drawn. For example, should I not buy clothes that I know might have been made by underpaid sweatshop workers, for this is treating them as a mere means towards my own aims to be fashionable? Such questions are still debated today, with the second formulation used to justify different positions on key issues of applied ethics.

**Activity**



Identify whether humanity (the agent) is being used as a means or an end in the following scenarios.

i) A scout helps a frail old lady across the road to earn his ‘Assisting the Elderly Badge’.

ii) A student steals his friend’s assignment so that he can copy his work.

iii) On the way home from the pub, Matt gives a homeless man one pound and thirty three pence.

iv) Belinda compliments her line manager’s terrible outfits and laughs at his bad jokes every day because she wants to be promoted.

v) A woman steals bread from the supermarket to feed her starving family.

vi) Orhan sleeps with Lucia to make his ex-girlfriend jealous.

vii) Steve buys everyone in the office a Milky Bar from the corner shop.

Strengths of Kantian Ethics

What are the advantages of Kantian ethics? Already you might have started to get an idea of where critics might have pointed out the flaws in Kant’s arguments, but it is important to note why a deontological ethical position such as Kant’s holds an intuitive appeal, and why Kant believed ideas such as the categorical imperative could prove an effective guide to moral action.

**Objective and Universal –** There are no ifs and buts in Kant’s moral theory. It presents moral laws that apply universally and impartially, so there are no cases when a person’s self-interest is routinely allowed to subvert the principles it puts forward, compared to utilitarianism, where people’s happiness is constantly weighted and judged. In this way there is an attractiveness to the idea that no one is above the moral law.

**Reasoned** – The centrality of reason to Kant’s ethical theory is often seen as an attractive element, for not only does it reinforce the potential for it to be objective and universal, but it also means that anyone who possesses a good degree of reason can grasp the moral laws it presents and understand their validity in different situations.

**Emphasis on Human Dignity** – We’ve seen in the previous section on utilitarianism that consequence-based ethical theories can have a difficult time accommodating the idea of human rights. But Kant’s second formulation of the categorical imperative establishes a strong foundation for natural or human rights in stipulating that no human being should be treated as a means, only an end. This is a very significant and important element of Kant’s ethics, and one that is often still discussed today.

**Emphasis on Autonomy** – Kant’s ethics is postulated on the basis of human beings have free will and autonomy. This for many people is intuitively important, and the emphasis Kant places on having the right intentions potentially highlights an important aspect of morality: the need for people to not simply perform the right action, but also perform it for the right reasons. Furthermore, there is not one defined end for all action, such as happiness. People can pursue whatever they wish so long as it is in accordance with the moral law, and informal duties such as the requirement to develop one’s talents potentially endorse individuals working towards ideals exemplified in their personality and ability.

Issues with Kantian Deontological Ethics

1. Are Consequences Really so Unimportant? – The Case of the Inquiring Murderer

**Imagine that one night you are woken in your bed by the sound of loud banging at the front door.** When you go to answer, standing before you is a vast figure who wields a machete and whose face is covered by a bloodstained ice hockey mask. He demands to know the exact whereabouts of a woman who you happen to know full well is spending the weekend at a hotel down the road. You even know her room number.

Now the sensible thing to do in this situation, if you could get your wits about you, would be to lie; to tell the murderer that you have never heard of this woman or even give him a false address. The one thing that would seem to be a terrible idea in this scenario would be to tell the truth: ‘Of course, she is staying at the Crystal Lake Inn, room 73A.’ Yet, for Kant, this would be the correct course of action because the maxim, ‘Lie when it suits you to do so’ cannot be universalised (if everyone lied whenever they felt like it, people would quickly stop believing anything anyone said).

A scenario similar to this was presented to Kant in his own time by one of the (many) critics of his moral philosophy. It has come to be known as the case of the inquiring murderer. Kant’s own response to the problem is widely considered unsatisfactory. Despite our moral intuitions to lie to the murderer, Kant suggests that we should still tell the truth, as if one lies then one is violating the entire basis and premise of the moral law to not lie. This, if universalised, would make society unworkable, and so Kant, rather than admitting a situation where the categorical imperative seems to fall down, bites the bullet and still argues that one has a duty not to lie, even to the murderer. In this way, one is only acting as to what is good by the moral law, not what the consequences of lying might be.

Many have criticised Kant here, and, although a few have defended his line of thought on a legal basis, it can still be argued that there is simply a direct conflict of moral intuitions behind Kant’s response. So what can be said in support of Kantian ethics in this situation? One line is to argue that a person is only responsible for their actions. If I lie to the murderer I am denying their rationality, treating them as a means to an end. In this way I am obliged to treat even the murderer as an end, and it is not my responsibility to restrict the freedom of the murderer, even if I think he is going to kill an innocent person.

However, James Rachels in *The Elements of Moral Philosophy* points out two problems with this counterargument:

(i) Firstly, is it really the case that our predictive powers are so limited? Determining the future behaviour of other human beings is arguably essential to a functional society. Supermarkets will stock more charcoal, burger buns and disposable cutlery in the summertime because they are confident that more people will be having BBQs in July than they will in January. Likewise, there will be more police in a city centre on a Friday or Saturday night than on a Tuesday afternoon because the police are able to predict that there will be more trouble from wayward drinkers at the weekend when the clubs are open late than on a weekday when the clubs are shut and the clubbers working.

(ii) Secondly, Rachels points out that ‘Kant seems to assume that although we would be morally responsible for any bad consequences of lying, we would not be similarly responsible for any bad consequences of telling the truth’ (p. 125). The issue here is that Kant is having his cake and eating it too; on the one hand, he wants us to be morally responsible for the consequences of breaking absolute rules (e.g. if the murderer found the victim inadvertently), on the other, he does not want us to be morally responsible for the consequences of following them (e.g. by telling the truth, we have arguably aided the murderer in finding and killing the victim). Either consequences are morally relevant, or they are not. Kant cannot have it both ways.

How else might we approach Kant’s response? Another response is to suggest that when using the categorical imperative, in the case of the inquiring murderer, Kant simply has not narrowed the conditions enough for it to be valid. Therefore, we might be able to instead formulate a universal rule such as ‘One should lie in order to save an innocent life’. However, this potentially introduces more difficulties, as how do we really know who is innocent? What if the murderer was actually seeking, unbeknownst to us, an individual who was instrumental in ending the lives of many Jewish people during the Holocaust? Introducing conditions not only makes us have to analyse a situation based on certain ends, but also complicates the moral decision-making process, and relies less on reason to guide our actions.

A similar response might be to argue that the second form of the categorical imperative restricts the requirement to not to lie to the murderer, as allowing them to go and kill another person is in effect treating the victim as a mere means towards the killer’s ultimate goal of murdering them. Yet such a response seems to present even more difficulties for Kant, for each of the formulations is supposed to simply be a reorganising of the one categorical imperative. If they were truly the same in essence then they should surely agree in their conclusions on what is the moral law in any situation?

It is possible, therefore, to argue that the case of the inquiring murderer presents a real problem for Kant’s ethics, and highlights three issues in particular:

1. It potentially shows that, at least in certain situations, consequences do matter, despite Kant arguing that an action’s moral worth lies in the intention and act itself.

2. It perhaps indicates that the categorical imperative does not provide real guidance in a moral situation. For, while I know I do have a duty to not lie, do I have a duty to tell the truth? Can I just not reply to the murderer? Can I tell a half lie, such as the hotel the victim is in, but not the room number?

3. Is there any room for supererogatory (beyond the call of duty) actions in Kant’s ethics? It might be the case that while not lying is generally the correct moral action, we might wish to praise the individual who lies to the murderer and risks their own life in the process. If all morality is simply following the law, is there any way we can praise those who seem to do more than the law requires?

#### Discussion:

*Do you think the case of the inquiring murderer is an unsolvable difficulty in Kant’s ethics? How would   
you go about finding a solution?*

### 

2. What Happens when Duties Conflict?

#### *Examples of Conflicting Duties*

Thinking about the case of the inquiring murderer earlier, we can recall that Kant believed human beings had an imperfect duty to care for other people. Yet while one is not obliged to act on an imperfect duty all the time, it seems strange that the perfect duty not to lie would clash with this imperfect duty so harshly. If we imagine that the inquiring murderer came every week to our door, asking for a new address, at what point might we be obliged to act on the imperfect duty to care for others and lie about the new victim’s whereabouts?

What proves difficult for Kant, however, is not just when imperfect and perfect duties clash, but when two directly universalisable duties clash. For Kant’s ethics depends on duties being consistently able to be universalised, and if there are other duties that clash and cause issues with a particular imagined duty, then it cannot be said that the duty in question can be conceivably universalised without contradiction.

For example, imagine in the case of the inquiring murderer that I made a promise to the victim I would not let the murderer harm them. If the murderer then approached the door, I would be forced to either lie or let the murderer harm them, breaking my promise. Kant might argue that I should not have made the promise to the friend, as I should have known I could not lie to the murderer if they approached me, yet, at the same time, it seems transparently obvious that I could have helped the victim by simply lying and so preventing harm to them. What this shows is that there ultimately is some difficulty in moral decision-making in Kantian ethics that altogether perhaps exposes some weaknesses in the usefulness of the categorical imperative as a whole.

#### Decision-making in Kantian Ethics

We can imagine lots of different scenarios in particular where imperfect duties conflict, especially as there are no direct guidelines for when they should be adhered to and when not. For example, although Kant states I have an imperfect duty in effect to give charitably, when does this apply, to whom and for what purpose? There are no substantial material facts included in Kant’s ethics, nor is there even context included for the kind of society a person lives in, and the traditions and practices it embodies. While Kant seeks to develop a purely rational account of morality, it can be argued that our moral decision-making necessarily incorporates information from the world around us, and not just on a purely conceptual level.

For example, say I have a parent who is in the later stages of severe dementia and cannot care for themselves. However, on the way from work to their house, where I plan to care for them for the night, I witness a person being hit by a car with no one around to help. I know I should help the victim and take them to a hospital ideally, but doing so risks the health of my parent, who, if left alone too long, can become disoriented at best and severely ill at worst. At this point there is an immediate duty of care versus a continual duty of care, but Kant’s imperfect duties do not capture the various options open in the situation, and if I sit there and try to calculate every option according to the categorical imperative I am likely to help no one. Should I help the victim of the car accident call an ambulance, and then help my parent? Or should I make sure that the victim receives all the help they need until they are in the full care of medical professionals, no matter how long it takes?

Ultimately, a strong criticism of Kant would be that the categorical imperative simply isn’t capable of resolving such dilemmas. They are context-sensitive, incorporate huge numbers of variables and relations, and involve numerous potential conflicting duties. A weaker criticism may be that there is a definitive answer in Kantian ethics, but it is far from immediately available. So, if I cannot even use the Kantian ethical system in ordinary life, what use is it in guiding moral action altogether?

***Conflicts between the First and Second Forms***

One contentious issue is whether there might be conflicts between the different forms of the categorical imperative. Such conflicts would be difficult for Kant to resolve, as it is supposed to be the case that there is only one categorical imperative simply emphasising the different ways one can reasonably arrive at moral laws. Yet there is a certain ambiguity in the way Kant approaches treating people as ends, which can lead to potential contradiction.

The most basic example might be seen in the case of the inquiring murderer. While for Kant we are obliged to not lie to the murderer (which would involve treating him as a means), we can still ask whether we are surely then treating the victim as a means to the murderer’s violent wishes? In fact, it can be argued that we are complicit in the murder at that point, meaning we are treating a person as a means to our insistence on upholding the moral law in the face of all existential circumstances.

Kant might respond that we are not obliged to look at the consequences of our actions in that way. Each act is relating only to the situation at hand, and we do not have to think of the secondary implications of our actions when conceiving of the moral laws we are obliged to follow. Yet such a restriction would seem strange to many. Instinctively there is an idea that we are aware of how we are treating not just the person immediately in front of us, but also the individuals that person is due to interact with. Therefore, it can be argued that once again Kant has difficulty resolving the conflicts raised when decision-making with the categorical imperative.



**Stretch and Challenge**

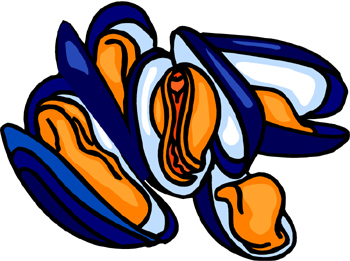
3. The Problem of Overly Specific Maxims

The obvious solution to Kant’s conundrum would be to make the maxims we act upon more flexible. Instead of following the maxim, ‘You ought never to lie’, why not say, ‘You ought never to lie *unless* doing so will save a person’s life’? There is surely nothing self-defeating about that maxim, and it would resolve the difficulties presented by the case of the inquiring murderer in a stroke.

The only trouble with this response is that it is unclear where we draw the line. It may become possible to make maxims so specific that any kind of action becomes permissible. Say my friend – call him Tom Reid – wants to borrow some money for a new PlayStation but has no intention of ever paying it back. The maxim ‘Take out a loan if you want it but don’t pay it back’ cannot pass the universalisation test because it is self-defeating; if nobody paid loans back, nobody would give out loans in the first place. However, the maxim ‘You ought to always repay loans unless your name is Tom Reid and you need a new PlayStation’ refers to such a specific set of circumstances that it is unlikely to result in any contradiction were it universalised. Yet we are unlikely to agree we have, thereby, made Tom’s action morally right.

Kant probably foresaw the flaws of this kind of reasoning. Crucial to his deontological ethics is the notion of consistency. If morality is to be binding, it must apply to all people equally. We cannot have rules that work for one but not the other. As Rachels puts it, ‘if you accept any considerations as reasons in one case, you must also accept them as reasons in other cases.’ Tom’s reason for allowing himself to borrow money he never intends to repay may be something like ‘Oh man, I really need that new PlayStation or everyone will think I’m a real flat tyre!’ However, if the shoe was on the other foot and it was Tom who was giving out the loan, would he really be willing to accept somebody else’s desire for a video games console and fear of being unpopular as a reason for them never to pay him back? It’s highly unlikely. One of the strengths of Kant’s position is that it ensures no one person is ‘special’ or has interests which are of greater significance than anyone else’s.

#### 4. The Issue of Trivial and Immoral Maxims

Alasdair MacIntyre claims in his book *After Virtue* (which will be explored more in the Virtue Ethics section) that the categorical imperative allows both trivial and immoral maxims to pass, as there is no actual moral content within the formula itself. He gives as examples:

**Trivial Maxim:** ‘Always eat mussels on Monday in March’

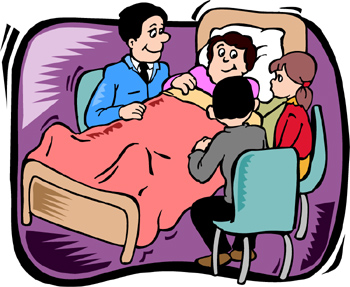
**Immoral Maxim**: ‘Persecute all those who hold false religious beliefs’ / ‘Keep all your promises throughout your entire life except one’

MacIntyre claims all these pass Kant’s conceivability test, as imagining a world in which they are true is not really contradictory. Taking the promises example, we can recall the original criterion by which Kant argued it would be inconceivable to imagine a world in which everyone had free rein to make false promises. Such a system would destroy the idea of a promise in the first place. Yet in MacIntyre’s example, in the knowledge that each person would only ever make one false promise we might regard it as perfectly rational to still trust promises. In fact, we could even make a law such as ‘Keep ninety five percent of your promises’ and still regard it as able to be universalised.

Similarly, in the case of persecuting those who hold false religious beliefs, it is arguable that it is not contradictory or inconceivable to imagine a world where this might be the case. Many people would put forward it is important to hold the right kind of religious beliefs, and so there is nothing necessarily immoral about pursuing those who insist on holding false ones. Critics of MacIntyre might argue here that he is mistaking the point of the categorical imperative, which is to determine which laws we are importantly bound by within the context of ordinary human life, not to simply generate whatever laws seem to make sense, but it can be contended that MacIntyre raises an important issue within Kantian ethics. In being divorced from empathetic ideas of morality, one can reason a number of different moral laws that seem either to be meaninglessly trivial, or run counter to our ordinary moral intuitions. This suggests that morality has a basis not just in reason, but perhaps also in a variety of different sources and ideas.

5. Is Duty the Only Morally Significant Motive?

Another common criticism of Kant’s ethics is his insistence on a dispassionate approach to morality. There is something rather cold about all good having to be done as it is what reason dictates, rather than out of love or the goodness of one’s heart.

The simplest example of this is given by Michael Stocker in his 1967 essay ‘The Schizophrenia of Moral Ethical Theories’. Imagine you are in the hospital after a severe accident and slowly recovering. During this time your friend comes to visit you, and, excited to see them, you remark ‘Thanks for coming, it’s nice to know someone cares about me’. However, your friend then turns round and says ‘I came to visit you not out of friendship but because it was my moral duty’. We might be perplexed by this, even angry, for it seems to devalue not only the meaningfulness of your friendship, but the entire idea of developing a closer form of care for human beings beyond one’s moral duty.

Looking back, we can recall that Kant requires that for an action to be good the motive *must* be to do your duty. This means, strictly speaking, that if you act out of love for a significant other or family member then your action has no moral worth. However, Kant can argue that he does not require us to feel *nothing* when we act morally, just that our choice should not be based on *feeling* but on *duty*. Only then can we make a rational decision, and act in accordance with the categorical imperative. So we might help our family out of love *and* because it is our duty.

Yet Stocker argues that this still misses the point. For if my friend came to see me in the hospital out of duty the moral worth of their action is still the same as if they came out of friendship also. At this point it can be literally stated that although under Kantian ethics it is okay to act out of feelings, those feelings still mean nothing. This is a strange prospect if one expands on it in the context of an entire friendship. For example, imagine instead that you are approaching the end of your life, and your best friend once again comes to visit you in the hospital, as they have always done when you have been ill. Except this time they inform you that they never cared for you, they were only acting out of their duty as a friend. This would be profoundly alien, and more than disconcerting. It can be put forward further that it is not enough for Kant to simply allow motives within his moral theory, he has to instead account for why these motives seem morally important beyond their conforming to the moral law.

We can also look here to the twentieth-century British philosopher Bernard Williams, who made an influential criticism of this line of thought. He asked us to imagine a situation where a man can save only one of two people from drowning in a shipwreck. One of them is his wife, the other a complete stranger. He knows that whoever he does not save will die. Of course, presuming the marriage is healthy, the man is going to choose to save his wife. This would likely be acceptable to Kant; a maxim stating that ‘You ought to save your spouse from drowning’ would likely pass the universalisability test.

Nevertheless, Williams points out that we have a peculiar arrangement where according to deontological ethics the man must justify his action as follows ‘I love my wife so I will save her *and* the maxim I am acting under accords with the categorical imperative’. The man has ‘one thought too many’. It seems absurd that we should need to be motivated by love *and* duty to do good.

(Note that this criticism also applies to utilitarian theories. Mill might allow a rule which lets people save their spouses from drowning because, on the whole, marriages bring people happiness and their dissolution brings them pain. However, for an action to be morally right do we need to do it out of love *and* to satisfy the principle of utility?)

In Kant’s defence, he never intended the categorical imperative to be a tool for practical decision-making. That’s why it occurs in a book titled *Groundwork of a Metaphysics of Morals*; Kant is trying to work out, theoretically, what morality is, and what it requires of us. He isn’t trying to do appliedor practical ethics. So it may be fair for him to say that in cases such as the above we are not required to think in terms of the categorical imperative. Whether such a response would be acceptable to Williams is for you to decide.

#### Is an Impersonal Perspective on Morality Possible?

These criticisms add up to a commonly made criticism of Kant; that his form of ethics simply does not align with our ordinary moral intuitions and the ways we apply morality in our lives. While his intentions are admirable in wanting to find the objective basis from which we might understand and develop a universal system of ethics, the result is instead an understanding that morality as a whole cannot be understood as a pure project of reason. Furthermore, it may simply be the case that human beings cannot adhere to a system of ethics that makes all moral decision-making wholly impersonal. It was explored in the Utilitarianism section how that moral system potentially violates personal integrity, and it is possible to push the same criticism against Kant’s deontological ethics. In the case of the inquiring murderer, we are not allowed to follow our own reasons for moral action and instead are supposed to submit ourselves to a wider project of reason, no matter how much it clashes with our moral intuitions. In fact, Bernard Williams makes the further argument that in practical moral deliberation this impersonal perspective is simply not possible; we cannot abstract away our personal reasons for performing an action to an impersonal one in order to achieve a universal unity of interest. In this way Kant’s morality requires us to abandon a notion of ourselves as moral agents, and simply become a vehicle for abstract reason.

However, in the next section we will explore a theory of normative ethics that arguably seeks to preserve such ideas of integrity by focusing not on moral acts, but on moral ethics. This moral theory places central importance on the idea of virtue, and how concepts of virtue might naturally push us to act in moral ways.

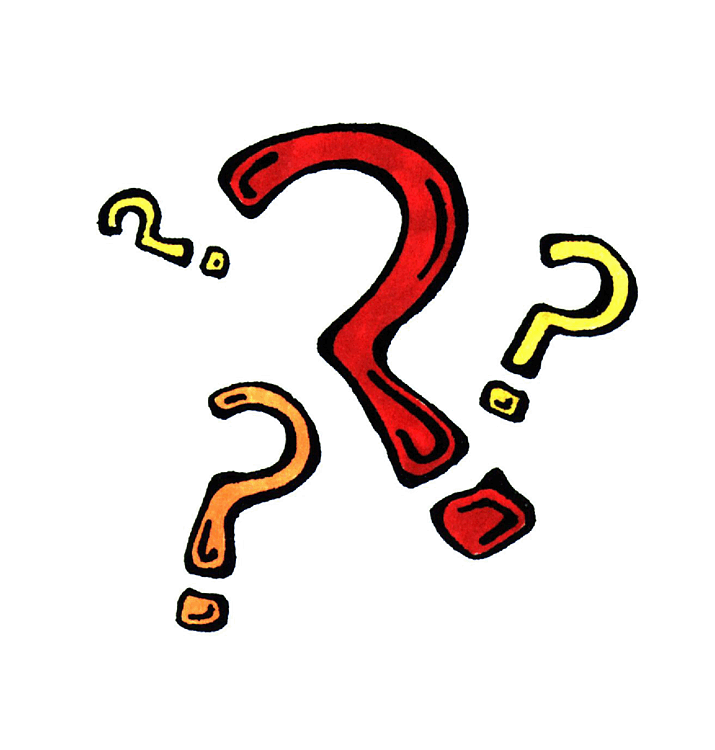
6. Philippa Foot – Hypothetical Imperatives as Foundation of Morality

Expanding on the issues of impersonality in Kantian ethics, Foot makes the more direct claim that morality cannot be understood as a system of categorical imperatives, only hypothetical imperatives. Without a particular end to our moral action, there can never be enough of a justification or reason to perform it as we are required to act independently of any motive or reason for the action itself. More succinctly we can ask of Kant, what reason do we actually have to obey the ‘oughts’ of the categorical imperative?

This is a thornier question than it might appear at first. Foot identifies that we do have non-hypothetical oughts in life, such as the ‘ought’ that one ought not to openly split infinitives in grammar. Yet it is difficult to reconcile these with moral action. While independent of a fulfilment towards a desirable or graspable end, these do not really carry a ‘command’ in the way that moral laws require. One can still split infinitives if one wishes as the grammar rule is only binding in a conventional sense, not a universal moral sense.

There is, therefore, perhaps an ‘illusion’ in Kant’s ethics that people are simply bound by morality and ought to follow moral laws for their own sake, out of respect for what duty and morality is altogether, but it is hard to argue that this convention is really binding. The person who acts immorally is not necessarily acting irrationally for Foot, they are simply not following the convention of respecting moral oughts for their own sakes. In this way a person can be consistent and immoral; they are not thinking about some conceivable world where their actions might be contradictory, they are focused on their actions in relation to a particular hypothetical end.

Foot, therefore, argues that a proper morality has to be built out of hypothetical ends, but ones directed towards proper moral goals that do not have self-interest at their heart. In this way there is no universal set of duties, but, as explored before, it may be a mistake of Kant to suppose that there needs to be one. Morality instead depends on you having the right moral motivations directed towards the right moral ends, and the question of whether an action is moral is not decided by an appeal to universal reason, but to a practical reason as to whether it fits within the correct moral motivations and ends.



**Kantian Deontological Ethics: Quick Quiz**

1. What is meant by deontology?

2. Give an example of a hypothetical imperative.

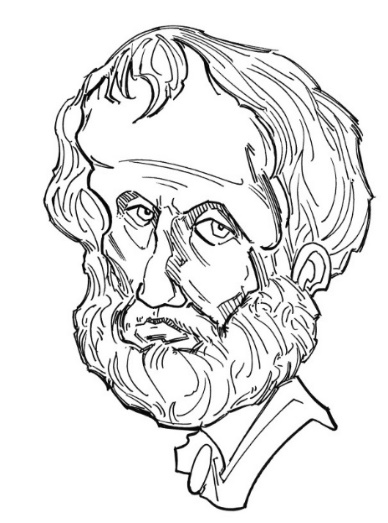
3. What is a maxim?

4. Give an example of a categorical imperative.

5. State the second formulation of the categorical imperative.

Ethical Theories 3: Aristotle’s Virtue Ethics

Introduction to Virtue Ethics

Virtue ethics as a field of ethical enquiry is one that is hardly new. Rooted in the works of Aristotle, particularly the *Nicomachean Ethics*, it was a common system of normative ethics in ancient Greece, and influenced greatly many philosophers and theologians, from St Thomas Aquinas to David Hume. Yet at the same time, for a large part of modern history it was not regarded as a system of normative ethics that could be taken alone. For example, virtue was often thought of as important in Enlightenment philosophy, but it was something to be fitted into a deontological or utilitarian ethical theory, where the failure to appreciate virtue was an argument against a particular ethical theory, but not an endorsement of virtue ethics itself. Similarly, in Christian theology, Aquinas gave a detailed account of the important of virtues but emphasised its importance within a framework of natural law, and worship of God.

It has only been in the last century that virtue ethics, directly inspired by Aristotle’s writings, has begun to emerge as a system of normative ethics by itself, and much has been written about whether virtue ethics as a whole is capable of effectively guiding moral action. Some philosophers view virtue ethics as an addition, or secondary project, to a utilitarian or deontological ethics, while others argue that the failings of other ethical systems point towards virtue ethics being a theory capable of standing by itself. Throughout this section, therefore, it is worth considering not only whether virtue ethics is capable of informing human moral action, but whether it is best understood as being contained within another utilitarian or deontological system of ethics.

**Anthology Text:**

Aristotle, *Nicomachean Ethics*: Books 1, 2, 3, 5, 6, 7, 10



#### Agent-centred and Act-centred Morality

One of the primary differences between virtue ethics, and deontological and utilitarian ethics, is that it focuses on the character and state of a person committing a particular moral act rather than the act itself. This means that virtue ethics is often called an agent-centred rather than an act-centred moral theory (although be careful as some philosophers reject this distinction). Therefore, when we look at whether a particular act was good in virtue ethics, we are not necessarily looking at what feature of the act made it good, such as whether it conformed to a particular rule, or maximised utility. Rather we are analysing whether an act fits a particular right, or virtuous, type of character. This is based on the intuition that when we think about moral people, we do not just look at their acts in the present. Rather we judge them and their act’s moral worth from their character, in the past, present and future, and how a particular act chimes or aligns with the kind of good character we wish to see in people.

Therefore, the questions asked by virtue ethics are often thought to be broader than those of utilitarianism and deontological ethics. It is not just the present act that is thought about; the virtue ethicist will also be questioning what makes a good person as a whole, and the best ways one can try to develop into a good person. To take an example, Christians in an ethical dilemma will not just read the Bible for answers, they will think ‘What would Jesus do?’ This is in the belief that Jesus was a perfect example of good character and virtue. If one can imagine Jesus performing a certain action, then that action is likely to be moral, even if it may lead to bad consequences, or break some particular rule.

In this way virtue ethics might be termed a more ‘holistic’ moral theory than utilitarianism or deontological ethics. It takes a wide variety of important environmental, natural and human facts about morality and the world, and seeks to develop a system of ethics that incorporates how we think about moral people, not just moral acts.

Aristotle and the *Nicomachean Ethics*

The best place to begin when studying virtue is with Aristotle, the philosopher who, it can be argued, pioneered the field of what is now termed virtue ethics. Aristotle has been greatly influential in many areas of philosophy, from metaphysics to epistemology, even though we only possess what many think are a fraction of his overall writings, with even those surviving thought to be more like lecture notes than completed works.

His main surviving work regarding ethics is the *Nicomachean Ethics*, which he puts forward as an empirical investigation into the way human beings naturally behave, and the best way to develop an ethical system to encourage the best qualities human beings possess. This is quite a different approach from his contemporaries, such as Plato, who regarded sense experience as ultimately being untrustworthy. Plato believed that only pure reason could discover what was moral. This means that moral laws or ideals could not be accurately reached through empirical investigation of the world. Rather, true knowledge of what was moral could be arrived at through reasoning on the nature of the Form of the Good.

But Aristotle thought such an approach was too abstract, and not grounded in the everyday facts of the world. Rather, what was important was the correct application of practical reasoning and wisdom; reasoning that helps us understand what to do, not just what is. Aristotle, therefore, contended that the kind of abstract speculation Plato was engaging in could not really help human beings become better people, and it was instead necessary to draw a system of ethics out of ordinary observations of human behaviour, and the goals such behaviour could feasibly be directed towards. Therefore, Aristotle’s ethics (and, generally, virtue ethics as a whole) is teleological. The correct forms of human behaviour are sought in order that human beings might reach a particular good end through such behaviour. This end of goodness, towards which virtuous behaviour should be directed, Aristotle termed ‘Eudaimonia’.

1: What is ‘the Good’?

Each ethical theory we have looked at so far has a specific idea of what ‘goodness’ consists in. For Kant, it was obeying the dictates of reason, for utilitarians, it was the maximisation of utility. For Aristotle, it is something called ‘Eudaimonia’.

There is no precise translation for *Eudaimonia* in English, but it may be taken as ‘good living’ or ‘flourishing’, and for Aristotle it is the goal, or telos, of all human action. But there are two things that can be questioned here. The first is to what extent is human action directed towards some good, and the second is why that good is believed by Aristotle to be Eudaimonia.

First, we can distinguish between doing something for its own sake (an end in itself), and doing something for the sake of something else (a means to an end). In the case of the utilitarian, they would argue all human action is a means to pleasure or happiness, as explored in the idea of psychological hedonism. Similarly, on a basic level, we can look at many of our actions and see how they are directed towards particular ends. For example, I might look to work and earn money so I can go the theme park. The aim of working is not simply to work in itself, but as a means to engage in my favourite hobby of going to theme parks.

Here we might readily agree with the first part of Aristotle’s argument. It does seem as if a lot of our actions are performed with some end in sight, although we might contest that certain actions, such as making art, are done for their own sake. Leaving such criticisms aside, however, we can still ask, why Eudaimonia?

#### Eudaimonia as the Final End of Human Action

Many philosophers have argued that defining Eudaimonia as happiness is leaving something out of Aristotle’s understanding of the concept. Were we to simply argue that the final end is happiness it would seem that Aristotle agrees with the utilitarian that happiness is the only thing that matters. Yet it is patently clear from his writing that Aristotle has a broader idea in mind when discussing Eudaimonia (hence why it is generally now translated as ‘flourishing’).

]

One important aspect is that it is seen as the final end of all human action. For sure, happiness is an important part of Eudaimonia, but it is feasible for Aristotle that a person could seemingly be happy but not have achieved the good life. In particular, Eudaimonia refers to the life well lived, where one cannot add anything that would make it more complete or better lived. This might be seen in a breakdown of the word itself with the prefix ‘eu’ meaning ‘good’ and the suffix ‘daimon’ meaning ‘spirit’. The whole word in a sense, therefore, refers to a spirit of life, a flourishing of the good beyond the ordinary conditions of life and luxury.

It might be contended, however, that Aristotle doesn’t really define Eudaimonia properly here. In fact, one might say that he has not provided a real reason to suppose that Eudaimonia should be thought of as the final end of action more than any other concept such as happiness, pleasure, money or just goodness itself. Part of the response is that it is impossible to define Eudaimonia in exact terms. Not only does it depend on the conditions of one’s life, but also it is a state which exists in relation to life, not as an object that can easily be grasped or understood. The other part of the response comes from how Aristotle begins to develop practical measures for guiding human beings towards Eudaimonia, and as part of this he considers why other ends of human action cannot in themselves be considered the final end.

#### Discussion:

*Both utilitarianism and virtue ethics are teleological ethical theories. Why do you think, then, that   
Aristotle and Bentham arrive at different ideas of the ‘end’ of human ethical action? Do you think pleasure or Eudaimonia is a more suitable moral aim?*

#### Why not Pleasure, or Goodness as the Final End?

Aristotle considers many different potential ends of human action, each of which he considers unsatisfactory compared to Eudaimonia. The two most important to consider are pleasure and goodness, which will be examined in more detail below. However, one can also consider other ends, such as honour and contemplation, which will be examined in a stretch and challenge section.

#### Pleasure

When we looked at utilitarianism, it was noted that one potential criticism of Bentham was that he failed to differentiate between higher and lower pleasures, as Mill did. However, it can even be argued that Mill’s idea of a competent judge of pleasures fails to prevent the possibility that human beings simply judge a life of base pleasure more important than one of art, philosophy and spirituality. In fact, Aristotle makes this criticism of pleasure, or more broadly psychological hedonism, arguing a life simply dedicated towards pleasure is no better than that of a cow or beast.

To judge from the lives that men lead, most men, and men of the most vulgar type, seem (not without some ground) to identify the good, or happiness, with pleasure; which is the reason why they love the life of enjoyment.

Therefore, Aristotle seems to be making the more blunt case that if one is dedicated to pleasure, one cannot help but live a vulgar life, as one fails to appreciate the importance of other goals and virtues. We noted how Mill argued even these goals could be thought to have arisen from their aims towards human happiness, but Aristotle dismisses any such claim from the start. As will be explored later, pleasure is important within a framework of virtue ethics, but cannot be thought to be the final end.

#### Goodness

The second important end Aristotle considers is that of simply goodness itself. Such an end would be closer to Kant’s vision of the good will or duty, where one simply acts well for the good itself and nothing else. Yet Aristotle holds a more environmentally, and perhaps realistic, sense of the good, for someone could act virtuously and be good all their life yet have a life of intense suffering due to some unpredictable physical condition. A person in this state of suffering, no matter how virtuous, could hardly be said to be flourishing. For example, later in Chapter 13 of the *Nicomachean Ethics*, when discussing pleasure, Aristotle states:

Those who say that the victim on the rack or the man who falls into great misfortunes is happy if he is good, are, whether they mean to or not, talking nonsense.

Similarly, he notes under his discussion on honour that one could feasibly be virtuous while being permanently inactive, even asleep. Yet, to take a more common example, we could not watch the Disney film of *Sleeping Beauty* and argue that the title character, despite seemingly being virtuous, is flourishing while permanently lifeless.

This might seem trivial at first, but it raises an important factor in Aristotle’s philosophy; that morality is not separate from the external, physical conditions of one’s life. As will be seen, being virtuous is a balancing act, often predicated on good fortune, and physical conditions outside of human control can greatly threaten a flourishing and good life.

***Aristotle’s Rejection of Honour and Wealth as Ends***

Aristotle does consider a few other ends of human action, rejecting each one in turn. It is not essential   
to know these, but they are included below for your interest.

Wealth – It’s a common theme even now that money should be the end of human effort. But Aristotle, similarly to many people now, argues that it cannot be the marker of a good life, and at best is simply a means towards things that make us happy, or lead us to flourish.

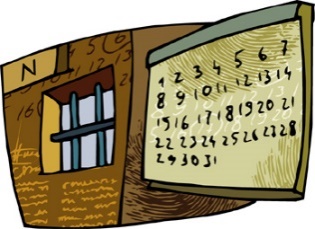
Honour – This is a more elusive end, and may even seem strange in our modern-day culture. However, in Greek society, there was much less of what we might call a ‘private life’ among ordinary people, and a big mark of worth was how a person was perceived among their peers. In this way, it would not be uncommon in Aristotle’s time for people to strive to be honourable. Nevertheless, Aristotle argues it cannot be a suitable end of moral action, as honour requires a person bestowing it, and people can change their minds, or be ignorant or fickle in their opinions about it. In comparison, Eudaimonia is something people understand themselves, and so isn’t based upon the whims of others.



**Stretch and Challenge**

2: The Function Argument

So now we have explored what Aristotle means when he talks about Eudaimonia, but the question can still be pressed as to how human beings achieve it. For one can easily understand how human beings might achieve pleasure (do pleasurable activities), but the question of how human beings flourish is a more difficult one to answer, especially since the idea of Eudaimonia itself is arguably difficult to pin down. In answer to this question, Aristotle gives what has come to be known as the ‘function’ argument; that is, to understand how human beings flourish, one has to understand how human beings function in the first place.

We will go through what this argument specifically claims, but, for the moment, consider the question, ‘What makes something good?’ For example, consider the items below, and consider what qualities would make them good rather than bad:

1. A rifle

2. An education

3. A prison

4. A person

For (1), we might easily put forward what makes a good rifle is its ability to shoot straight, its comfortability, and its reliability, among other qualities. For (2) we might bring in more broad qualities, such as its ability to prepare someone for a job, or the way it conveys important information. For (3) it might be its ability to keep prisoners incarcerated and how effectively it manages to rehabilitate them on release.

Yet (4) is what Aristotle is most interested in. For he argues that all things have a function, directed towards an end, and what makes something good is how well it performs its function towards that end. Therefore, a rifle that shoots straight can be called good as it possesses what Aristotle terms the necessary *arête* (virtue/excellence). A rifle that makes any bullet miss a target, no matter how trained the shooter, is one that is bad. So when Aristotle turns to the question of what makes a person ‘good’, he is asking how well that person is achieving their function. For any person, therefore, we can examine them and determine to a certain extent how well they are excelling according to their functions and, therefore, determine whether they are on the path to fulfilling Eudaimonia.

Understanding this, we can now lay out the premises of Aristotle’s function argument:

1. The good for human beings is performing their functions well

2. The characteristic human function is ‘an activity of the soul which follows or implies a rational principle’

3. Therefore, the good for human beings is performing this activity of the soul which follows… a rational principle

#### Issues and Criticisms

1. The good for human beings is performing their functions well.

We examined previously how an object might be said to be good by how well it functions. However, such good is commonly described as ‘instrumental good’, and many have contended that instrumental good is not the same as moral goodness. In particular, when we looked at the naturalistic fallacy, and the is–ought gap in relation to utilitarianism, it was noted that perhaps there was a logical gap between identifying a natural end to human actions and equating that with what is good (see Meta-ethics section later for more information!).

So, in short, why should we believe Aristotle when he states that human beings fulfilling an instrumental good is the same as fulfilling the moral good? When Aristotle states that a certain activity of the soul correlates to what is good, we might simply disagree and say that very thing is actually bad. However, there are a number of ways of looking at this issue.

The first is to understand the context of Aristotle’s thought. The philosopher Alasdair MacIntyre, for example, contends that a system of ethics can only be effectively understood within the context of the society and culture it arose from. When we look back towards Aristotle’s era, and the society he lived in, it can be noted that there was the idea of each person having a specific role in ancient Greek society, whether it be farmers, soldiers, or even slaves, and these had associated qualities. For Aristotle, therefore, the idea of individual people having functions was not far-fetched. This notion was commonplace and so Aristotle accepted it as a commonly understood feature of reality rather than a contentious claim.

However, reducing Aristotle’s claim to history may not be doing it justice. While Aristotle appeals to the different occupations and roles people have in society, he also appeals to the natural ways human beings function. For example, we have eyes that function to see things, we have hands that function to grasp things and we have noses that function to smell things. If all these individual things have function, is it so difficult to imagine that human beings as a whole have a function?

Overall, Aristotle, therefore, makes an argument from analogy, comparing the parts of the body to the overall whole. Yet it is patently obvious that this analogy is weak. While human beings have eyes that have a function, this is very different from the entire human person, and moreover we can contend here that Aristotle is committing the **fallacy of composition**. This was observed in Mill’s proof for utilitarianism, and concerns when an assumption is made that the whole of something shares properties with its parts. In this way, just because the parts of a human body have a function does not necessarily mean that the whole human person has a function.

These criticisms prove difficult for Aristotle, for ultimately a key part of the function argument relies on human beings having distinct functions that it is good for them to fulfil. Therefore, if one undermines the teleological view that underlies the function argument, one arguably undermines the entirety of Aristotle’s ethics. However, efforts have been made to recover the basis of virtue ethics, as will be noted in the next section.

#### Discussion:

*Both Mill’s proof of utilitarianism and Aristotle’s function argument fall prey to the fallacy of composition.   
Is this a real issue for these arguments, or are they simply making an intuitive move from individual cases of function and happiness to a general, aggregate case?*

#### Virtue Ethics: Beyond Function

The Greek word Aristotle uses, that is often translated as function, is *ergon*. However, it can be contended that this word extends to more than just function, including some notion of work or characteristic, and it can be argued that when Aristotle asks what is the function of human beings, or contends that they have a function altogether, he is more asking what the defining characteristics and natural activities of human beings are. In this way, we can focus more on a broader understanding of function, that isn’t simply limited to what we might think human beings are good at.

One idea that might support such an idea is that of genetics, or genetic function. When the basis of human survival and reproduction is analysed, it can be argued that there are basic functions that underlie such efforts. These might include cooperation, comprehension and other activities that ensure in a way a survival and overall flourishing of the human race, such that when we extrapolate these natural activities into the moral sphere in which human beings operate, we end up with a set of virtues that promote these functions. Therefore, moving beyond the assumption of teleology Aristotle makes, one can look more empirically at the way human beings biologically function and how this might relate to ethics as a whole.

Alternatively, one might take a position such as that of MacIntyre, who takes elements of Aristotle’s teleology and incorporates them into a wider understanding of how human societies function. Rather than say human beings have a defining function or set of functions, he argues that what we consider to be moral goods have to be understood within the societal contexts and practices from which they emerge. Therefore, in different societies there may be many different goods corresponding to the kinds of roles and practices a particular society endorses, and each society’s goods cannot be necessarily compared with another’s as they arise from very different circumstances and practices.

There is great philosophical discussion ultimately both on whether human beings can ever be thought to have definitive functions, and, perhaps more importantly, where these functions arise. While Aristotle’s argument may seem initially logically weak, it can be simply viewed in the same breath as Mill’s proof of utilitarianism, an argument that relies on a number of intuitive or common-sense propositions about how human beings operate, and how this might relate to what we consider to be the moral good.

2. The characteristic human function is ‘an activity of the soul which follows or implies a rational principle’.

If we provisionally accept the first part of the function argument, we can then ask, what are the functions that are characteristic of a human being? Aristotle approaches this question by examining the natural world, and noting the differences between things, species and beings, cancelling out those characteristics which human beings share with other beings. Through this process he believes he can identify the primary *ergon* of human beings.

However, it is important to note that this process ties in with Aristotle’s view of the soul. Whereas his contemporaries, such as Plato, often viewed the soul as an immaterial, eternal aspect of a human being that persisted beyond death, Aristotle took a much more scientific view, identifying the soul of a being as its set of characteristics and the form of those characteristics. In this way, the soul of a plant were the parts or characteristics that constituted a plant, and defined it as separate to other living things. In this view you may well see the underpinnings of modern biology and how it defines life, and it is true that much of modern biological science is arguably built off Aristotle’s understanding of the natural world!

But turning towards the question of the human soul, and its defining characteristics and functions, a deeper question emerges as to what separates us from animals and plants. Aristotle contends it cannot be something such as growth, for that characteristic is shared by all beings. Neither can it be a perceptual or sentient drive, for animals also possess that characteristic. Instead, the one unique aspect of human beings that Aristotle identifies is rationality, and the way human beings act upon their life with rational principles. This distinct characteristic, therefore, forms the basis for Aristotle’s famous definition of the function of human beings.

The function of man is an activity of the soul which follows or implies a rational principle.

In this way, the function of human beings is not simply to act according to our sentient appetites or desires, but to reason upon them, and evaluate our actions according to the rational principles we can discern as being important. Yet, it can still be asked, what is the relation of this activity of the soul to virtue?

#### Virtues, Reason and the Soul

Turning back to the concept of virtue once more, we can examine closer the Greek term *arête* that Aristotle uses throughout the *Nicomachean* *Ethics*. While the word ‘virtue’ has many different connotations and meanings now, for Aristotle, as noted previously in the function argument, it simply meant the characteristics that allowed a thing to be classified as good or excellent. In this way, it is sometimes better to translate *arête* as ‘excellence’, and when the term ‘virtue’ is used, think about it more in terms of the way ‘virtuosity’ is used to describe someone who is particularly skilled at a particular practice.

When we talk about virtue, or *arête*, in this way, it becomes clearer how it relates to the function of human beings. For it has been noted that the good, or Eudaimonia, comes from human beings excelling in those functions unique to human beings, that is the rational activity of the soul. Therefore, virtues for human beings are simply those characteristics and qualities which allow human beings to excel in this function.

It is in passage 7 of Book 1 that Aristotle completes this line of argument, stating:

… human good turns out to be activity of soul in accordance with virtue, and if there are more than one virtue, in accordance with the best and most complete.

So, for Aristotle, living by the virtues which help us exercise our rational activity of the soul is what leads us to Eudaimonia. To achieve the good, one has to maximally fulfil one’s function, but in order to maximally fulfil one’s function, one has to excel in the right kind of virtues. However, defining these virtues, and living by them, is arguably the more difficult part, to which most of the *Nicomachean Ethics* is directed.

This is hinted at the end of Book 1, where Aristotle notes how Eudaimonia fits in with our intuitive understandings of happiness and flourishing. For example, we would not say that after learning our first times tables we had mastered mathematics, and, in the same way, mastering the virtues is something that takes place across our entire lifetime. There is not one simple state of being non-virtuous, and another of being virtuous. Furthermore, Eudaimonia is not something that is achieved through our rational activity alone. External conditions will affect our ability to excel in our functions, whether it be war, disease or hunger, and achieving the virtues is not simply a matter of finding some inner ability to reason, but rather something produced by the cultivation of habit. All these aspects will be explored in the next section.

The twentieth-century French philosopher Jean-Paul Sartre claimed that ‘existence precedes   
essence’. This was intended as a direct rebuke to thinkers such as Aristotle who assumed the opposite:   
that each thing has a particular purpose that is somehow essential to it. For Sartre, there are no grand plans or imbued functions – particularly for human beings. In fact ‘man is thrown into the world’, meaning that objectively life is quite meaningless and it is up to the individual alone to bring purpose to it.



**Stretch and Challenge**

3: Virtues and the Doctrine of the Mean

#### What is a Virtue?

Before we continue on to discussing how Aristotle believes the virtues can be determined and applied, it is important to say a little bit more about what a virtue is. We’ve noted how it is a characteristic or trait that allows something to excel in its function; however, it is important to understand that for Aristotle this is not something necessary as part of a particular thing, nor is it represented in a single act. For example, we might say that a person was kind for giving a homeless person money, but we might stop short of saying they are a kind person altogether, for we do not know how they act outside of us seeing that single act of kindness.

Therefore, a virtue is a characteristic that represents a propensity for something to reliably act or behave in a particular way. Our rifle we used as an example earlier has the virtue of being accurate if time after time it shoots straight, not just if it happens to shoot straight once. Furthermore, with wear and tear the rifle may become worn down and cease to shoot straight, in which case it loses its virtue of being accurate. Nevertheless, these are just singular examples of virtue; how more broadly in modern philosophical terms can we understand such characteristics?

#### Discussion:

*Aristotle wrote* Nicomachean Ethics *from the perspective of a culture vastly different from ours today.   
Do you think his ideas of what ‘virtue’ is still have relevance today? Can you intuitively understand what might be virtuous and what might be vice-like in your own life?*

#### ‘Thick’ Ethical Concepts

Virtues in moral philosophy are often thought to hold a somewhat unique place in ethical discussion. Thinking back to utilitarianism and deontological ethics, it can be noted that they revolve around discussions of what is ‘good’ and what the term ‘good’ might reasonably be applied to. However, when examining virtue we are not just examining what is good, but we also seem to talking about a wider set of descriptive claims that refer to the function and characteristics of a particular thing.

In this way, virtues are often referred to as a ‘thick’ ethical concept, as rather than just talking about the prescriptive, moral aspects of a particular thing or action, it contains key descriptive elements about it as well. For example, if I were to call my friend untrustworthy, I would likely be inferring that they have a morally bad character. But tied up in this moral judgement is a number of descriptive claims about the person also. For me to call them untrustworthy means that in the past they have regularly and purposely said things which are lies, and have some deficiency in their ability to tell the truth at certain key times. Furthermore, I am implying that this deficiency is not temporary. From their lying in the past I expect that they will lie in the future, making a predictive claim about their character. Therefore, altogether, in calling my friend untrustworthy, I am making a number of important descriptive claims about their behaviour, and then using such descriptive claims to make a moral judgement about them.

Such ‘thick’ moral judgements have been argued by philosophers to be important when discussing virtue ethics in not only normative ethics, but also in the context of meta-ethics, where critics of naturalistic philosophies have claimed that one cannot go from descriptive claims about the world to the normative claims of ethics. ‘Thick’ ethical concepts, in providing both descriptive and prescriptive elements in their use, are potentially one possible way of bridging the is–ought gap in moral philosophy.

However, more importantly within the context of Aristotle’s ethics, you should be able to understand how the idea of virtue is multifaceted and complex. Achieving virtue is not just a matter of performing a certain act, but cultivating traits and characteristics that enable one to reliably and purposely do what is virtuous in any situation. How this is done is covered extensively by Aristotle in the *Nicomachean Ethics*, and one important concept, which will be covered next, is habit, which arguably is the crucial link between performing virtuous acts and becoming virtuous.

#### The Importance of Habituation

At the beginning of Book 2 of the *Nicomachean Ethics*, Aristotle states:

… it is also plain that none of the moral virtues arises in us by nature; for nothing that exists by nature can form a habit contrary to its nature. Neither by nature … nor contrary to nature do the virtues arise in us; rather we are adapted by nature to receive them, and are made perfect by habit.

What does Aristotle mean by this? One good place to start is by thinking about virtue in relation to your own life. For example, as a baby you were unlikely to be patient. If you were hungry you would have cried and not considered that occasionally you might have to wait longer for food. Growing up as a young child, it is also likely that often you would not have been patient either. For while your parents might have withheld things from you at necessary times, you might not have understood the importance of this and similarly been impatient with them. Yet growing even older, you might begin to recognise the importance of patience and begin actively cultivating it in your life, such that in any given situation you display patience beyond when it is seemingly socially required.

In this way, through the education and cultivation you become patient through habit, and Aristotle emphasises the importance of habit particularly in the development of virtue. For, while human beings possess the potential in their natures to develop the virtues, it is only through habit that they are properly developed. One cannot simply reason on what is virtuous, and expect to become virtuous. Rather it is through understanding the virtues, and regularly practising what is virtuous that one effectively becomes virtuous altogether. In this way, when Aristotle talks of habit, he is talking more about habituation, the reasoned process of learned repetition, the same as when we begin learning maths or any other skill throughout our lives.

#### Emotions and Virtue

As we have seen, Aristotle talks about developing the virtues as if one were developing a particular skill. In the same way as when we begin learning a musical instrument, the early stages of developing a virtue are difficult and often painful. There is a steep learning curve as one is introduced to a new skill and the habits that accompany it. However, after enough practice and habituation, the practising of the skill (or virtue) becomes more pleasurable, in the same way that a musical virtuoso playing their preferred instrumental derives pleasure. Furthermore, there is the move from being educated in a skill, to being the one educating others in one’s knowledge of it.

Yet we can note some differences between the cultivation of a skill such as learning an instrument, and the cultivation of a virtue. For one, the development of a normal skill might involve pleasure and pain, but it does not readily involve as much moral significance. One might get angry and frustrated being unable to play a bit of music, but many would say anger might be more acute in the face of injustice. In this way the development of a virtue, on the other hand, has deep moral significance for the way we live our lives. Encountering those who readily lie for their own ends, while being expected to be correctly honest ourselves, can produce deep-seated emotional responses as one is forced to deal with moral unfairness, and the burden of seemingly taking the high road in all moral decisions.

Aristotle, therefore, emphasises virtue as not just the right response to a particular reasoned or technical challenge, but also to our emotions. The virtues cannot be equated with certain emotions, but have to be formed alongside them, and with respect to how emotions, and their responses of pleasure and pain, drive our lives. An individual emotionally conflicted about whether or not to give money to a homeless person in this way likely has not become a generous person, for the cultivation of this virtue has not developed to the point where their response is habitually generous. Therefore, it is important in the process of developing virtue that one also develops the right kind of emotional and pleasurable response to being virtuous. If one lets negative emotional states continually fester when attempting to be virtuous, then the process of acting virtuously is much more likely to be interrupted and fail. So habituation is a process involving both rational reflection and emotional development. For Aristotle, one without the other cannot lead to a completely holistic cultivation of virtue in a person. This is especially important when we turn to one of his most famous ideas, the doctrine of the mean.

#### The Doctrine of the Mean

To work out what can be considered a virtue, Aristotle said moral agents need to find the mean or **balance between two vices**, avoiding excess or a deficiency of any virtue. This is what is known now as the doctrine of the mean, and is a primary starting point for how Aristotle believes human beings can come to understand virtue and excel in their own lives. He writes:

Virtue is a mean because the vices respectively fall short of or exceed what is right in both passions and actions, while virtue both finds and chooses that which is intermediate.[[1]](#footnote-1)

To work this out, individuals need to use **practical wisdom (*phronesis*)**. This is an individual process of working out what virtue is suited to each situation and to each individual. Thus there can be no rules or maxims about how to act. Individuals use their autonomy and intellect to work it out and then put it into practice.

Aristotle says that man is able to work out for himself what is good – not just what will produce what is good in a particular situation, but what will produce the good life in general:

Now it is thought to be the mark of a man of practical wisdom to be able to deliberate well about what is good and expedient for himself, not in some particular respect, e.g. about what sorts of thing conduce to health or to strength, but about what sorts of thing conduce to the good life in general.[[2]](#footnote-2)

We can take the virtue of courage below as an example:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Vice** | **🡨** | **Virtue** | **🡪** | **Vice** |
| *Example:* | *Foolhardiness* | *Courage* | *Cowardice* |

The alternative to a virtue is a ‘vice’ in Aristotle’s thought, yet, as with virtue, this term now carries a number of unhelpful connotations. We might refer to smoking cigarettes as a vice but this does not come close to what Aristotle means when he uses the Greek term *kaika* to refer to them. For a vice nowadays is often just something we consider to be bad, but for Aristotle it refers more to a defectiveness of character, the opposite of virtue as excellence.

So, considering courage, we can say a few things about how it relates to its vices foolhardiness and cowardice. Say, for example, you and a friend are travelling through the forest, and all of a sudden a bear jumps out behind a tree and starts mauling your friend; what would your response be?

Well, if it were to just run away and leave your friend for dead, no one would be that impressed, especially Aristotle, who would argue that in running away you committed an act of cowardice. On the other hand, if the first thing you do is to jump on top of the bear’s back and begin punching it, one would be inclined to say such an act is just foolhardy, for it won’t do much damage to the bear, and won’t really help your friend. All that will likely happen is both of you get hurt.

So what might be the mean to these actions? Well, one interpretation of Aristotle’s teaching has been to suppose that it refers to some moderation between two vices, yet such an interpretation is not only unhelpful, but profoundly wrong. While moderation may be important in some situations, it doesn’t really make sense to apply such a principle here, where moderation might mean you only kick the bear once and then slowly run away. Really what Aristotle refers to is perhaps better expressed as a balance between one’s emotional responses and rational choices.



In the case of the bear attack you are likely to feel a number of different conflicting emotional states. Part of you will be scared, which will want to motivate you to run away, while part of you, perhaps fuelled with adrenaline, will want to save your friend. Yet being spurred on only by this emotion will obscure any rational considerations of the situation and cause you to do something rash or foolhardy. So, it seems, Aristotle wants us in the doctrine of the mean to mediate between these two emotions (and the vices they motivate) and, along with using our practical wisdom, discover a path in between them, or, as he states in Book 2, an ‘intermediate between excess and deficiency’.

So when the bear attacks, you might pause for a moment to think about the different choices and emotions facing you, before deciding the most courageous thing to do is shout for help while looking through your bag for the bear mace you packed, using it to defend your friend against the bear.

Nevertheless, at this point you might state, well how am I supposed to know to do that from just using my practical wisdom and considering my emotions? Well, for Aristotle, you’re not. If you remember cultivating virtue is something that takes a long time, through education and habituation. If the bear attack was your first time encountering danger without education or supervision it is likely you would choose the wrong course of action. It is only perhaps through encountering similar situations requiring similar acts of courage, along with a more experienced park ranger’s education about bear attacks, that you might be able to act effectively. As Aristotle states in Book 2:

For in everything it is no easy task to find the middle … any one can get angry- that is easy- or give or spend money; but to do this to the right person, to the right extent, at the right time, with the right motive, and in the right way, that is not for every one, nor is it easy; wherefore goodness is both rare and laudable and noble.

This might seem strange; surely moral choices should be able to be intuitively understood, otherwise what is the point of developing an ethical system in the first place? But we can say Aristotle here is simply being realistic. One would not expect a baker to operate as effectively on the battlefield as a soldier; the latter has been trained not only in the best rational way to respond to dangerous situations, but has had experience of courageous acts in battle that allows him to naturally understand what is courageous on a battlefield. There may be times when the soldier still chooses a path of vice, and it is not guaranteed that education and reason always lead to virtue in every situation, but the habituation developed through life as a soldier inevitably helps the soldier, if he has the will, to make courageous decisions more effectively.

However, there is lots to unpack here, and we will explore a few criticisms of Aristotle’s doctrine of the mean before we progress on to moral responsibility within Aristotle’s ethics.

#### The Virtues in Full

|  |  |  |
| --- | --- | --- |
| VICE OF EXCESS | MEAN | VICE OF DEFICIENCY |
| Rashness | Courage | Cowardice |
| Licentiousness/Self-indulgence [recklessness; libertinism; a waster] | Temperance [self-restraint] | Insensibility [unable to enjoy anything; ‘insensibility is not human.’ NE, Bk. 3, Ch. 11] |
| Prodigality [reckless spending] | Liberality [generosity] | Illiberality/Meanness [miserliness] |
| Vulgarity/Tastelessness | Magnificence | Pettiness |
| Vanity | Magnanimity [being a ‘good sort’] | Pusillanimity [weak-willed; craven] |
| Ambition / Empty vanity [arrogance] | Proper ambition / Pride | Unambitiousness / Undue humility [a lack of self-respect] |
| Irascibility [short-tempered] | Patience / Good temper | Lack of spirit / Unirascibility [disinterested; bland] |
| Boastfulness | Truthfulness | Understatement / Mock modesty [bashfulness] |
| Buffoonery | Wittiness | Boorishness [dullness; uncouthness] |
| Obsequiousness [a door mat] | Friendliness | Cantankerousness [moodiness] |
| Shyness | Modesty | Shamelessness |
| Envy | Righteous indignation | Malicious enjoyment / Spitefulness [Schadenfreude] |

*Based on p. 104 of the translation by J A K Tomson*; *bracketed items are explanatory comments.*

#### Issues: Is It Always Morally Appropriate to Identify Virtue as a Mean between Two Vices?

This is a question that often initially emerges once beginning to study the doctrine of the mean. Taking our previous example of courage, it would seem that cowardice is a much more extreme opposite of that virtue than rashness is. Indeed, in many situations it is arguably hard to distinguish what is courageous from what is rash until after the action itself. A soldier running through gunfire to save a wounded friend may simply die, even though his act arguably is courageous, and without access to certain inner mental states we might be hard-pressed to not regard their action as a little rash. In fact, Aristotle directly addresses this issue at the end of Book 2:

To the mean in some cases the deficiency, in some the excess is more opposed; e.g. it is not rashness, which is an excess, but cowardice, which is a deficiency, that is more opposed to courage… This happens from two reasons, one being drawn from the thing itself; for because one extreme is nearer and liker to the intermediate, we oppose not this but rather its contrary to the intermediate.

What he states is that from the nature of excesses and deficiencies in the virtues we are studying, one vice often seems more contrary to the intermediate than the other. This is reflected in both the nature of the vice itself, and the way we approach it. Some things simply are more contrary to the mean, but this does not reflect on to what extent the virtue is an intermediate, just that in one case a particular vice as commonly experienced in the world is further away from the intermediate than another. For example, self-indulgence is much more common than insensibility, and those who do self-indulge often push it to greater limits than those who restrain themselves. This means that when we think of what might be the opposite of temperance, we naturally think of self-indulgence rather than insensibility.

#### Are the Vices Listed Always Vices?

This is quite a broad question that will be explored further later, but we can ask to what extent Aristotle’s idea of vice is culturally determined. For example, while he views unambitiousness as a deficiency, another view might be that even proper ambition is generally harmful. Similarly, we can ask whether modesty or truthfulness are really virtues when in many cases they seem to hamper discussion and action by human beings. We might envision a time when being shameless is generally good, or lying is more conducive to a good state of affairs. The basis of these issues will be revisited, but as a starting point it is worth working through Aristotle’s list of virtues, and questioning where you view the vices as potential virtues depending on the situation, culture and context.

#### Difficulties with the Doctrine of the Mean and Human Emotions

The last issue we can briefly look at is whether when using the doctrine of the mean, we should really be required to mediate between emotions, as well as rational considerations of deficiency and excess. For example, say I discover the vicar of the local church has been secretly stealing money from townspeople and using it furnish a lavish villa in the South of Spain. Should I not be angry at this, and should my anger not motivate me towards right action against this example of hypocrisy and injustice? We might classify this as righteous indignation, but what difference are we really supposing there is between this virtue and the vice of irascibility, for I certainly shouldn’t be patient in my condemnation of the vicar. In fact, Aristotle admits that he struggles to find corresponding excesses or deficiencies for certain virtues (e.g. proper ambition) and some contend that rather than Aristotle’s system of virtues being universal, he has simply mapped what he considers to be good characteristics onto the framework of the doctrine of the mean. To what extent, therefore, one can divide emotions, and their corresponding vices, into camps of excess or deficiency is debatable, and it may be that looking through the table of virtues, you might question whether Aristotle has correctly assigned certain emotional states to real vices and virtues.

4: Moral Responsibility

#### Discussion:

*Is the doctrine of the mean essential to virtue ethics? How else might we discern virtue if not between   
two vices?*

We can now move on to Book 3 of the *Nicomachean Ethics*, where Aristotle deals with the question of moral responsibility. Arguably this is a more difficult to understand philosophical concept in virtue ethics. In utilitarianism, we can simply align moral responsibility to the degree to which someone should foresee the consequences of their actions, and in deontological ethics the extent to which someone should be aware of the correct moral law. Yet, as we have seen, Aristotle himself argues that virtue is hard to achieve, and not purely dependent on one’s own rational choices, so to what extent can we argue someone is really responsible for their virtues and vices, and the actions that result from them at any given point?

The first thing Aristotle states is that people should only be held responsible for **voluntary** acts. But what counts as a voluntary action? Most simply it can be understood as when we act according to our intentions; when we decide upon an act as the result of a choice based on deliberation. If we had the intention to perform a certain act, but always acted in another way as a result of a microchip in our brains directing us, then Aristotle would contend our action isn’t really voluntary. As a result it isn’t really an indication of what virtues and vices we have and so we cannot really be held responsible for it, for if we could follow our intentions then we would have chosen otherwise. This seems fairly intuitive; we wouldn’t regard external causes of our actions as real indications of our moral state, and many would think that our internal moral intentions should count towards being praiseworthy or blameworthy for our actions. But issues arise when we begin to consider the reasons why Aristotle believes that in certain cases we should not be held morally responsible for our actions.

#### Involuntary Acts

Aristotle first turns to the case of involuntary actions: those which are done under direct force or compulsion. The particular example he gives of these forms of actions is a person being taken somewhere by kidnappers or men in power. In this case there is no intention the person being kidnapped could have had, and nothing they could have done, in which case we cannot really hold them morally responsible; they could not have feasibly done otherwise no matter what their intentions were.

However, the second example he gives illustrates the more problematic form of involuntary acts. Imagine a sailor transporting some goods for his merchant bosses, and, upon his ship hitting a storm, he decides to throw the goods overboard to ensure that he arrives safely at his destination. Now, it cannot be said the sailor’s intention was always to throw the goods overboard; he was prompted by the arrival of the storm and fears about his safety. But, at the same time, we can argue the sailor did make a choice in the moment, and chose the priority of certain survival over the prosperity of himself and his bosses. So in a sense the sailor’s action was both voluntary and involuntary; how does Aristotle reconcile these two parts of the act? Should we accept the sailor was compelled to throw the cargo over, or should he be held responsible for his actions?

Aristotle argues that although there was a degree of compulsion in the sailor’s situation, they should still be held morally responsible. In particular he states:

… the man acts voluntarily; for the principle that moves the instrumental parts of the body in such actions is in him, and the things of which the moving principle is in a man himself are in his power to do or not to do.

In this sense, because the sailor does have a real choice whether or not to throw the goods overboard, we have to respect this choice as voluntary, even if in the abstract and long term it would be unwanted. For if, for example, the sailor had not thrown the goods overboard and delivered them safely, we would have praised him as courageous, and in similar situations where we had been courageous we would also want to be praised as courageous. In fact, we might agree with Aristotle here that such difficult situations are the ones which give rise to the opportunity to be virtuous, and without them we could not exercise the excellences we had habitually trained within ourselves. If we were so desperate to avoid blame in difficult situations then we should equally be desperate to avoid praise, but no one really claims upon doing something good or virtuous that they were forced to do it.

However, what Aristotle’s argument here does allow is degrees of moral responsibility. While we might under Aristotelian ethics blame the sailor for his actions, we can admit that he was compelled by factors outside of his control, and show understanding in our condemnation. This is quite important, as it gives room to distinguish cases where one can hold someone responsible for the purposes of moral improvement, and holding someone responsible for the intentionally bad acts they performed. In the former, we might accept someone is good, but was compelled towards vice by particular external factors and adjust any punishments accordingly. Therefore, while there may be some confusion about moral responsibility and certain involuntary actions, it can equally be argued that Aristotle is simply acknowledging the difficulty of simply dividing people’s actions up as either voluntary or involuntary.

#### Non-voluntary Acts

The other cases where we might regard someone as not morally responsible for their actions can be classed as ‘non-voluntary’. This is not a term that can easily be directly translated from Aristotle, but in Book 3 he does spend time discussing actions that are not intended being performed from ignorance or a certain lack of knowledge. These kinds of acts cannot be said to have a direct force or compulsion performing them outside of the agent’s intentions, but they cannot be said to be willingly performed either as, if the agent had more knowledge about a particular situation, they may have chosen another course of action. Yet, at the same time, we would want to also blame the person who performed an action both out of ignorance and bad intentions. For example, read through the two situations below:

1. Julie wishes to teach her friend Anna how to dance, as Anna last week told her she always wanted to learn how but could never find someone to teach her. She invites Anna to her dance hall, and in preparation puts on some music and lights. However, she does not know that Anna is epileptic, and on arrival she begins having a fit on the floor.

2. Jill has a rivalry with her boxing partner Sarah, and even when sparring with her to improve her technique she tries to knock her out whenever possible. Normally she is prevented by the heavier 16 oz sparring gloves they wear, but one day, unbeknownst to her, there is a mix-up and both Jill and Sarah are given 10 oz gloves to spar with. Jill as usual still attempts to knock Sarah out and this time succeeds as the gloves are lighter, with Sarah falling to the floor and seriously injuring herself in the process.

Now, in example (1), we would be inclined to not want to blame Julie, especially if, after making sure Anna was OK, she told us, ‘I wouldn’t have put on the music and lights if I had known Anna was epileptic’. The bad consequences of her actions in example (1) seem contrary to her good intentions, and she was just ignorant about Anna’s condition.

But in example (2) we would still want to blame Jill, even if she did not know the day of knocking out Sarah that the sparring gloves were lighter than usual. In fact we may ask her if she regrets her actions after the sparring session and she might say ‘no, I’ve always wanted to knock her out, and it just so happened by mistake today’. In this case, the lack of repentance on Jill’s part means that her ignorance was irrelevant to her actions.

These two examples illustrate how Aristotle thinks about non-voluntary actions. He states:

Everything that is done by reason of ignorance is not voluntary; it is only what produces pain and repentance that is involuntary. For the man who has done something owing to ignorance, and feels not the least vexation at his action, has not acted voluntarily, since he did not know what he was doing, nor yet involuntarily, since he is not pained.

Therefore, everyone who acts out of ignorance and regrets their action Aristotle regards as having committed an involuntary action, and, although they should be held responsible, they also should likely be pardoned. However, the person who expresses no regret at their action out of ignorance, while still committing a non-voluntary act, should be considered morally responsible for their actions, for their intentions always aligned with the end result of their choice.

Altogether, for Aristotle it can be seen that moral responsibility is a difficult and complicated issue. A person may have many reasons for their action, some of which are influenced by factors outside of their control, and similarly it may often be the case that an individual is ignorant about the conditions and consequences of the action they have chosen. In this way, in analysing any ethical dilemma from an Aristotelian perspective, one has to not only seek an understanding of whether the action was virtuous or not, but also the various influences, factors and information that might have caused a person to act a certain way, before attributing praise or blame for their actions.

***Issues with Moral Praise***

One question that can be asked of virtue ethics is how we are supposed to measure moral praise in   
relation to moral effort. For example, should we value more the person who was born more virtuous,   
and able to act more consistently virtuous by their personality and nature, than the person who has to   
strive five times harder to cultivate good habits and become virtuous? We can imagine a person who is born and raised psychologically courageous so in a particular ethical dilemma does not have to suffer worries that come with fear, but should we regard this naturally courageous person as being morally praiseworthy if they have put no effort into being virtuous?

This is a difficult issue, and Aristotle bites the bullet here and argues that moral effort should not impact moral praiseworthiness in this situation. However, others disagree, and argue that moral effort should correspond to praise in the cultivation of virtues. Nevertheless, if one endorses this view, one also leads to the issue where a person might engage in vice-like behaviour for an extended period of time, before becoming virtuous. In this situation, despite their having previously willingly engaging in vice-like behaviour we would be required to give them more praise for managing to overcome their vices and become virtuous.



**Stretch and Challenge**

5: Practical Reasoning, Contemplation and Pleasure

In the final section, we will consider the more nuanced ways Aristotle outlines how human beings become virtuous. As has been examined so far, it is far from a simple process, with many different aspects playing into the cultivation of virtue, and the responsibility and blame people face in their virtuous or vice-like moral actions.

#### The Role of Reason

It has been stressed a lot how reason is central to the right functioning of human beings, and the way Aristotle distinguishes between theoretical and practical reasoning. However, more can be said about the relationship between reason and a person’s voluntary actions. For example, even though voluntary acts come as a result of a person’s decision-making process, Aristotle in Book 3 argues there are some situations where a person’s voluntary decisions should not be held up as an example of their character. In the case of children, who are not fully rational or capable of making decisions, it could not be fair to judge them for the actions they do, even if they are made through a similar decision-making process. Similarly, in spur-of-the-moment cases, where a person’s action may result more from such a brief operation of the mind, it might be hard to judge them on it against a wider understanding of their character.

These cases highlight an important aspect of Aristotle’s thought hinted at earlier; that voluntary acts are important in discerning the virtuous acts of a person when they involve deliberative choice. This is where a person rationally considers various alternatives and makes a choice as to what they believe to be the best course of action. It is through exercising deliberative reason that human beings can effectively build a virtuous character, and, although it has been emphasised through this guide that there are physical conditions which potentially influence a person’s development beyond their rational choice, this does not undermine the centrality of reason in a person cultivating virtue themselves.

For example, it is reason that allows a person to observe virtuous behaviour, understand what a virtuous person does and then choose to act in a similar way in their own lives. It is reason that allows a person to understand what might be the mean between two different vices, and it is reason that allows a person to realise how both the intellectual and moral virtues are essential in reaching Eudaimonia.

#### Developing Practical Wisdom

However, so far we have mainly focused on how a person develops the virtues of character, the non-rational part of a person, and, while Aristotle emphasises the importance of habit in cultivating virtue, he dedicates a large portion of Book 6 to discussing one key aspect of reason: practical wisdom. We can distinguish this once again from theoretical reasoning (covered in Book 10 of the *Nicomachean Ethics*) as the form of reason that governs how we decide to do things. Therefore, for Aristotle, when we are faced with a choice, both our non-rational desires and our practical rationality influence how we decide upon an action and then follow it through. An important part of cultivating virtue, and acting virtuously, is the correct use of practical wisdom in any ethical dilemma.

And let it be assumed that there are two parts which grasp a rational principle-one by which we contemplate the kind of things whose originative causes are invariable, and one by which we contemplate variable things; for where objects differ in kind the part of the soul answering to each of the two is different in kind, since it is in virtue of a certain likeness and kinship with their objects that they have the knowledge they have.

To a certain extent practical wisdom as an intellectual capacity is a difficult to define concept for Aristotle. As a process of practical reflection upon the conditions of the world, our own non-rational desires and the different ways we might act in a situation, what is appropriate and non-appropriate action is not something that is immediately accessible, else it would be able to be known more easily through theoretical reasoning. The emphasis on practical wisdom strongly differentiates Aristotle from his contemporary Plato, who thought what was good, and the actions that could follow it, could be known better by a pure application of theoretical reasoning.

Yet Aristotle does give a number of ways in which practical wisdom can be seen in human activity, and these ways give insight into how human beings can use their practical wisdom in everyday life. He contrasts it in particular with the idea of scientific knowledge, which he argues is built off induction. But practical wisdom is not a similar form of knowledge, but a process built on top of scientific knowledge which works to inform our practical reasoning. Therefore, practical wisdom is not something that can simply be learnt, but is a prudential faculty which uses our knowledge, intuition and insight to develop a course of right action. In particular, Aristotle identifies a number of ways this faculty is seen:

1. Human beings are able to think about not what is simply good for them, but what might be a general good. In the case of virtue, this is the difference between simply acting according to one’s desire, and acting according to what might be reflection of general virtue.

2. Human beings can effectively judge their actions, and understand what might be the right thing to do from an idea of good, but also how the material facts of the world relate to that good.

3. Human beings can understand the broader significance of their actions, and the way they impact the wider world.

4. Human beings are able to foresee the way their current actions might influence and form their future selves. For Aristotle this is key to understanding habit, and practical reasoning helps us realise why engaging in a certain set of behaviours might lead us unavoidably to act in a similar way in the future.

5. Human beings possess the ability to follow through on their decisions, and perform the necessary actions to ensure they are effective. Practical reasoning in this way is not just a process of reflection, but a continuous engagement with our decisions and the attempts to make sure they come about the way they were intended.

***Aristotle and Contemplation***

One of the more perplexing aspects of the *Nicomachean Ethics* comes in Book 10. Through his work,   
Aristotle outlines the importance of practical wisdom, but near the end he turns to a different principle   
and process: contemplation (*theoria*). This is the arena of theoretical reasoning, and, while undoubtedly   
important in discerning what could be right and virtuous, it has been argued that Aristotle’s conclusions and thoughts in Book 10 stand against his philosophy in the rest.

While Aristotle asks whether the good life could be considered to be contemplation in Book 1, he only gives a list of reasons why this must be the case after developing the case for practical reasoning, and there is conceivably a tension between Aristotle defining a life well lived as one of contemplation versus the very hands-on approach to ethics he emphasises through the *Nicomachean Ethics*.

For example, Aristotle praises contemplation, as it is the ultimate use of reason, the defining function of human activity, but it is difficult to see how the exercise of reason by itself fits in with the forms of practical wisdom outlined in solving ethical issues, and the way virtues fit between two vices. Similarly, he regards contemplation as being a process of the highest good, as it is engaged in without requiring any particular end; for its own sake. It can, therefore, bring happiness, as nothing is required to engage in it, and there is no material end to it as an action. It allows us to achieve real knowledge of the world, and its highest concepts. So altogether Aristotle seems to present a life of philosophical enquiry, not practical virtue, as the highest good, even though such a life could hardly be said to encourage general human flourishing, or even embody many of the virtues Aristotle presents earlier on. There is a tension between the individualistic life of contemplation, and the virtuous life given to the community, such that Aristotle’s thoughts by the end of the *Nicomachean Ethics* resemble those of Plato much more than when he began it.

While it is important to note that theoretical reasoning has an important part to play in discerning virtue, and the forms of life around us, Aristotle introduces a whole new idea here in it being the only good activity leading ultimately to Eudaimonia. This part of his book has, therefore, sometimes been thought to be underwritten. However, others have contended it is reconcilable with his wider thought, suggesting that engaging in contemplation is a central example of human well-being, but is not equivalent to it. In this way, a flourishing life might be focused around contemplation, but a wider variety of factors of ethical virtue play into and structure this life.



**Stretch and Challenge**

#### The Role of Pleasure

It was noted earlier in the guide that Aristotle rules out pleasure as the final end of human conduct, yet at the same time he also seeks to avoid the claims made by philosophers such as Plato that pleasure potentially got in the way of people doing good. Instead he seeks to develop a realistic account of pleasure, one which acknowledges how it can drive people to turn away from what is good, but also can motivate people towards good as a natural function or feeling people get in life. In this way pleasure is not an end to be sought in moral action, but an important passion to shape in the cultivation of virtue.

This can be seen in the cultivation of virtue through habit. If trying to develop a particular habit is always fraught with pain, then ultimately its development will be hampered. The natural human instinct is to turn away from what is painful, and, for Aristotle, it is pointless to deny that. Equally, excess indulgence will prevent someone from seeing the real value in something separately from the pleasure it produces. Therefore, when we are trying to cultivate positive habits, we want to take pleasure in this cultivation, such that we become positively inclined to repeat the habits in our lives. For example, if I were to seek to become more generous, I would try to find pleasure in giving to others, which would encourage me further to give more in the future. So long as I keep in mind the end goal of the virtue itself, and not the pleasure it produces, the habit formed will correspond with the state of becoming generous as a whole.

Aristotle can, therefore, be thought to be outlining a ‘middle way’ regarding pleasure. It is not something to be abhorred, but equally is not a final end of human conduct. Instead it is a valuable way to help us build a virtuous character when it is focused on and developed properly.

The Strengths of Virtue Ethics

We might trace out the practical steps of Aristotle’s virtue ethics below:

Contemplation on the Good

Knowledge of the World

Desires/Instincts

Material Circumstances

Practical Wisdom / Deliberation

Choices/Decisions

Voluntary Action

Habituation/Repetition

Virtue/Excellence

From here one’s virtue would equally reflect back on one’s contemplation of the good, knowledge of the world and desires. One might also view this as a circle, but, as will be observed in the Issues section, this opens up Aristotle to a number of criticisms. What is most important to note is that Aristotle’s ethics is far from an abstract enterprise. It is heavily involved in the material aspects of the world, and so, in a certain way, is messy and difficult. The path to becoming virtuous is a difficult one, and the steps above will not always be followed regularly or correctly. For example, while practical wisdom might lead us to correct moral choices, any of the previous factors might change and influence the way we progress towards virtue. We might contemplate there is another sort of good we haven’t considered, some new knowledge about the world might arise, we might realise a certain desire is holding us back or even just the material circumstances of our lives prevent us from effectively cultivating virtue at all. When analysing Aristotle, though, be careful in simply arguing he is wrong based on the ineffectiveness of the virtue ethics method, as a part of its content and structure is that morality is built upon a shifting platform, which we often can’t fully grasp.

However, it is still possible to trace out a number of key strengths of virtue ethics:

**Agent-centred** – Compared to utilitarian and deontological ethical theories, virtue ethics takes stock of the intuitive importance of understanding morality as relating to not just principles and rules, but the person performing the action as well. In fact, one might contend that virtue ethics is the only really personal ethical theory studied so far that respects how human beings make choices and how they view their integrity and character in relation to them.

**Respects human desires and passions** – Rather than being a wholesale endorsement or rejection of desire, and the non-rational elements of human life, virtue ethics arguably well charts a system of ethics that accommodates them within a reasonable framework of moral life. One might say, therefore, that it does not risk falling into the trap of seeming too aloof or impersonal in its descriptions of ethics as a whole.

**Complexity** – Virtue ethics respects that ethics and ethical decision-making are complex issues, and that there are not simple rules or principles behind human action that we should follow. This means ethical decisions might be made based on numerous factors, from personal integrity to general good, and it rejects that all human beings, in being immoral, are simply ignorant or failing to respect basic moral rules.

**Centrality of virtue** – While other ethical theories try to incorporate ideas of virtue, virtue ethics has the advantage of placing virtue at the centre of its conception of morality, and so captures an instinctive way of thinking about moral character and action that has persisted throughout history.

**Practicality** – While other ethical theories often simply declare the moral rules human beings should follow, virtue ethics attempts to provide also a descriptive account of how human beings can seek to become more moral, and the difficulties associated with this. In this way it also arguably presents a more holistic, psychological account of morality, as well as a normative one.

Issues with Aristotle’s Virtue Ethics

1: Can virtue ethics give any guidance on how to act?

The distinctive feature of virtue ethics is its agent-centred, holistic approach to morality. Yet this is also arguably one if its greatest weaknesses. By focusing entirely on the long-term cultivation of the moral character, virtue ethics ignores entirely the need for guidance in the fast-paced world of moral decision-making.

The natural response from the virtue ethicist is that the virtuous person does not need exact instructions on how to act because their highly refined character will automatically be able to determine the morally correct choice in any given situation. Is this really believable, however? It seems as though there are certain ethical dilemmas where choosing the right or wrong option is not so much a reflection of one’s ‘good’ character but rather a result of painstaking deliberation. We are told to be brave and not foolhardy or cowardly, but how do we know whether the situation calls for bravery (e.g. storming the enemy) or the exercise of another virtue (e.g. patience)? This might be an even bigger issue when thinking about large-scale ethical dilemmas. Can virtue ethics really give an answer on the best way to solve political and humanitarian crises, for example? It might be a failing of virtue ethics that it cannot address the moral big picture, only the small decisions in a person’s life.

Aristotle says that we will know which act is the virtuous one through the exercise of practical wisdom. But what if we don’t have practical wisdom? Can we then never be virtuous? And if we all have practical wisdom, how come people still give in to the vices?

This last point proposes a dilemma, but it is an oversimplification of the role of practical wisdom in the *Nicomachean Ethics*. Aristotle is clear that practical wisdom is not some rare talent where ‘you’ve either got it or you haven’t’ (an ethical X factor if you will), rather it is a skill that can be worked on or allowed to waste away. We all have the potential to do X (for example, to draw, to paint, to dance) but only those who work very hard practising X have any chance of attaining mastery over it. Likewise, those who neglect, or even scorn X, will likely lose whatever aptitude they once had for it. For the virtue ethicist, then, morality is not a series of rules or maxims to be followed, but a lifelong commitment to honing one’s behaviour.

***Discussion:***

*In small groups, try to think up some more scenarios where two or more virtues may conflict with one   
another.*

#### Hursthouse and v-rules

However, some modern philosophers argue that there are certain forms of rules contained in virtue ethics. Rosalind Hursthouse, for example, argues that the virtues give rise to what she calls v-rules that arise out of the normative element of Aristotle’s account of what is virtuous. For example, if one takes the virtue of benevolence or generosity, it can be argued that moral laws such as ‘one should give what one doesn’t need to those who do’ can be derived from a full understanding of what these virtues mean. This means that if one does not follow the v-rule given by a certain virtue, then one is falling short of the behaviour that virtue requires. These don’t necessarily have to be binding, but can be general principles of human conduct that one can use as a framework during habituation until one is sufficiently virtuous to not necessarily require the rules at all.

Nevertheless, there are still issues with this idea, most notably that there is no meta-framework that governs how we should derive v-rules from the different virtues that we can identify. For example, one society’s idea of friendliness might differ from another’s such that the potential v-rule ‘one should smile when greeting people’ might be incorrect depending on the context in which one applies the virtue. Another case might be ‘one should always tell the truth’ as a v-rule for honesty, but many people don’t necessarily associate honesty with always telling the truth, rather with telling the truth at times of moral significance. This means it is not always clear what v-rules could be derived, and a v-rule such as ‘one should be friendly’ or ‘one should be honest’ is vacuous as it uses the virtue as an example of action without elaborating anything new.

One final potential response to this from the virtue ethicist might be that appreciating v-rules is a matter of practical wisdom. There is no guarantee a rule is right in every context, only in some, and it is up to the person seeking to cultivate real virtue through v-rules to know when they can be applied. However, at this point it is difficult also to know what real status v-rules hold, if they can’t be taken to be actual rules applicable to anything more than a few precise circumstances. Therefore, it is still possible to push the objection that virtue ethics fails to provide guidance on how to morally act.

2: What if virtues clash?

According to Aristotle, a clash of virtues is not a possibility; if you are truly virtuous (i.e. have practical wisdom) then you will know the correct virtue to apply to any given situation. If you think that the virtues are in conflict, then you’re simply not looking at the issue correctly.

Yet is it not conceivable that virtues may clash? For instance, I may have the virtue of honesty, but find that this clashes with my friendliness when a close companion asks what I think of their favourite band, The Who. My honest opinion may be ‘They’re washed-up old men’ but I am aware how unfriendly this may seem. Aristotle also maintains that the virtuous individual should have ‘proper ambition’; but what if having the right ambition occasionally requires one to boast (to do well in an interview), to be obsequious (to get on the good side of a new boss), to be miserly (one may need to save for a new home or car to live up to some kind of dream)? Do we not then have a conflict? Unless Aristotle provides a hierarchy of virtues (he does not), this is a hard problem to resolve.

When faced with such difficulties, proponents of virtue ethics are likely to point to the flexibility and holism of their philosophy. As a result, time and circumstance can easily be accommodated; so it may be that fulfilling the virtue of right ambition requires some obsequiousness or miserliness. Yet the person who balances all these things will be exercising their practical wisdom well, and so can still be considered virtuous.

#### Discussion:

*How could you envision virtues clashing? Do you believe there is a way of reconciling these clashes, or   
does it show virtue cannot be the only factor in our moral decision-making?*

#### Hursthouse – moral remainders

The issue of virtue clash is not something novel to virtue ethics; in fact, we’ve observed how utilitarianism and deontological ethics have their own forms of internal clashes of rules and principles and Hursthouse argues that virtue ethics is better equipped to deal with such clashes. Why?

Well, one issue with utilitarianism, for example, is that it can seemingly endorse certain right actions even if they produce greatly bad consequences, due to them producing marginally better consequences overall. For example, if I was forced to kill a third of my family to save the other two thirds from becoming zombies, it may be for a utilitarian this is the right action, and I should be praised for my actions. But after the act I am unlikely to feel as if I should be praised, and most likely will be racked with guilt about my actions, even if I reasoned they were right. This aftermath is what Hursthouse calls the ‘moral remainder’, and she argues it is something not acknowledged properly in utilitarian or deontological ethics, as they focus on the act and/or its consequences, but not the effect upon the character of an individual.

But virtue ethics, Hursthouse argues, does accommodate this ‘moral remainder’; bad actions have a lasting effect on our ability to be virtuous, and, in this way, while in virtue ethics clashes of virtues are difficult to resolve, this is contained within its acknowledgement of the moral remainder. We might feel guilty for choosing honesty over friendliness in a situation, for example, and hurting a friend with the truth, but that is part of the function of practical wisdom. There might be a right virtue to follow in a particular situation, but the possibility of a moral remainder is always present. Human beings in virtue ethics live with the results of their actions, and this plays an active part on how we reason about, and cultivate virtue in, our lives.

Such a response arguably does not completely eliminate the problem of clashing virtues, but it does soften it, showing that the clashing of virtues is not necessarily a fault in the structure of virtue ethics, but in the way human beings perhaps have to make ethical decisions in an imperfect world.

3: Is Aristotle making a circular argument?

This is clearly circular: how can we know that a person is virtuous if we don’t already know what the virtuous acts are? How can we know what virtuous acts are if we don’t already know who the virtuous people are? It’s a ‘Which came first, the chicken or the egg?’ scenario.

One way out is to talk about virtuous people in terms of their character traits, which do not consist solely in action but also in the way a person thinks, what they find pleasure in, their inner emotional life and so on. However, many of these determinations of character are not visible to everyone else. In the A Level Philosophy of Mind topic, you will discuss the significance (or lack thereof) that these inner mental states have. For now, however, the key point is that we can only observe the actions a person makes (their behaviour), we cannot observe how they feel or the way they think. So we are back to square one: we can only learn about virtue by observing how virtuous people behave.

Is there any way round this? Well, Aristotle could argue that we can infer the way somebody thinks or feels from what they choose to do, or what they say, or what they write. Yet these are evidently actions, and it may be that there are circumstances where not acting exhibits virtue. I display my temperance by *not* getting blind drunk at a friend’s dinner party; I display my patience with a student by *not* throwing a chair at them; I show my friendliness by *not* laughing at a co-worker’s bizarre new haircut.

A stronger response may be to argue that we do not need to know what a *particular* virtuous person is like, but rather just have an idea of what a virtuous person is like *in* *general*. Then we would be able to know such and such an act is virtuous because that accords with my general idea of the virtuous person. This is fine as far as it goes, but it raises an important question: who decides what virtue is like *in general*? Some philosophers (for example, MacIntyre) here bite the bullet, saying that virtue can’t be understood outside of a relevant society, and its traditions and culture, while other virtue ethicists have argued that a systematic study of virtue in different cultures can potentially reveal core universal virtues that are present in all human cultures.

***The charge of elitism and the Nietzschean challenge***

One criticism that can be levelled at Aristotle and the virtue ethics tradition generally is that they are   
inevitably culturally relative. On this view, all Aristotle’s theory amounts to is a prolonged description   
of those personal characteristics that his culture (essentially the Ancient Greek nobility) found admirable   
or distasteful. Likewise, the modern virtue ethics of Anscombe and MacIntyre lists those personal characteristics that mid twentieth-century British philosophy professors find admirable or distasteful. (It’s worth noting that these philosophers were rebelling *against* relativism, which they felt was leading the world into something of a moral abyss.)

Some thinkers would take this further: all moral philosophy is culturally relative and, moreover, elitist. In Mill’s utilitarianism we find the ethical proclamations of an emotionally damaged Victorian colonialist; in Kant, the kind of universal morality that may suit enlightenment Prussia, but is irrelevant to the fractured world of today. It is also notable that since the time of Plato, philosophers have placed themselves on something of an ethical pedestal. It is always they, the philosophers, who know what the good is; it is they who are writing the great books and gesticulating wildly in the forum or the lecture theatre; it is they who know not just what is good for them, but good for you too.

This line of thought finds its most powerful expression in the German writer and philosopher Friedrich Nietzsche, who argued that one of the great mistakes of Western philosophy was to assume morality could be discussed in a vacuum. For Nietzsche, morality has a history. There is a story behind why one kind of person or action is considered good and another kind is considered bad. He laid out most clearly what he thought was the story of Western moral philosophy in a book titled *A Genealogy of Morality*.



**Stretch and Challenge**

We don’t need to know the details of Nietzsche’s story, but   
central to his account is the observation that whereas the Ancient Greeks praised courage, strength and   
conquest, his society, heavily influenced by Christian teachings, praised meekness, self-denial and   
frugality. If morality can differ so wildly through time and place, why should we trust the claims to   
ethical truth made by one thinker at one time in one place? Nietzsche’s bold   
conclusion is that any claim to objective morality is bound to fail. Indeed, he   
considered himself to be living in an age when the entire façade of ethics was   
beginning to collapse, and that its replacement would be a world without values.

In defence of virtue ethics (and normative ethics more generally), ideas should not be   
judged on their origins, but on their merit. To argue otherwise is to commit the **genetic   
fallacy**; a type of argumentative mistake where an idea is rejected out of hand because   
of who said it, or where it originated. Just because Aristotle praised those characteristics   
he personally found amenable does not mean those characteristics are not – objectively –   
good. The reasons *why* honesty or generosity came to be valued may differ from place to   
place, but they are valued nonetheless. Nietzsche gives us cause to radically question received   
opinion, but the baby need not be thrown out with the bathwater.



**Stretch and Challenge**



4. Are there virtues that do not contribute to Eudaimonia?

This is potentially a thorny question for Aristotle. His ethics relies on a certain connection between virtue and Eudaimonia, but what if we might imagine a situation where exercising certain behaviours might be regarded as the right thing to do, although not contribute to overall flourishing in Aristotle’s system of ethics?



For example, imagine the case of an allied prisoner of war in a German POW camp. They are, with the help of a number of their fellow prisoners, trying to dig a tunnel out of the camp to freedom so they can rejoin the war effort and defeat Hitler. To this end, every day they steal tools from their prison labour to keep digging the tunnel, lie to the guards about their actions when they inquire in their cell and display an overt lack of self-respect in their lives so the guards think they have crushed their spirits. Moreover, the tunnel takes a long time to dig, so they have to keep up this pretence over a very long time. Here we might argue that their excessive humility, dishonesty and thieving are far from good actions, but in this context they are virtuous, for we would not criticise the prisoner for engaging in these behaviours, and even admire their zeal to keep fighting in the war. Yet at the same time we can hardly say these virtues, which Aristotle would consider vices, could be argued to contribute to Aristotle’s idea of Eudaimonia. So what is happening here?

A strict interpretation of the issue would call back to the issue of action guidance in virtue ethics, and claim here that the central problem is that there is not a real link between virtue and right action. Often they will coincide, but virtue ethics fails to account for large and deep problems of ethics involving what we might ordinarily consider out duties and principles. In the case of the prisoners, all that is happening is that the idea of good character is tied with the recognising of important moral duties or principles that the prisoners are attempting to follow.

A more moderate interpretation might be seen in the work of David Hume, who acknowledges the importance of virtue, but argues it is not tied to anything such as Eudaimonia. Instead, the moral judgements we make of a virtuous person’s character are representative of some emotional state we hold towards that person. When someone does something we consider virtuous, we have a sympathetic feeling of approval, and when someone does something we consider a vice, we have a repulsed feeling of disapproval. In this way, virtues are related to what we find internally agreeable or disagreeable, not to an overall ideal of Eudaimonia.

However, a virtue ethicist might equally respond to this objection saying that once environmental factors are taken into account, it is possible to argue that the prisoner’s behaviour is not virtuous. Rather it is the very conditions of their imprisonment that undermine the possibility of virtue at all, and prevent any real flourishing occurring. For example, say the prisoners did escape through a tunnel just as the war had ended, and suddenly they found themselves back in their home country, but having been used to lying and thieving while imprisoned, they found it hard to shake these behaviours and got sent to domestic jail as a result. The point may be that, although their actions may have been right in context (Aristotle even admits that occasionally a non-virtuous action may be right from practical reasoning), the very conditions of their imprisonment make cultivation of real virtue impossible. Therefore, from bad luck they are forced to cultivate vice, and this is simply an example of how difficult it truly is to become virtuous, and why education and the correct environment are important to uphold in any society or culture.

Nevertheless, some may find such a response still unsatisfactory, and argue it is simply dodging the inability of virtue ethics to explain why the prisoner’s behaviour is widely considered right despite contradicting Aristotle’s assertions.

***Is Aristotle’s virtue ethics overly individualistic?***

This is a subtle objection, but one which has often been charged against Aristotle’s ethical system.   
Bertrand Russell, for example, in his analysis of the *Nicomachean Ethics* argued that the kind of great,   
virtuous person Aristotle outlines is also the person many would feel instinctively uncomfortable with. For   
if we met someone wholly ambitious, magnanimous, liberal and righteous, we might regard them as a bit self-involved, especially if all their moral thinking revolved on their personal quest for virtue and not in order to help those in need.

Indeed it is a point Russell stresses that the *Nicomachean Ethics* is light on any morality centred around the needs of others, to the point where we might argue that virtue ethics as a whole in its individualism goes against our moral intuitions. For do we really conceive of morality as simply a path towards our own happiness and flourishing?

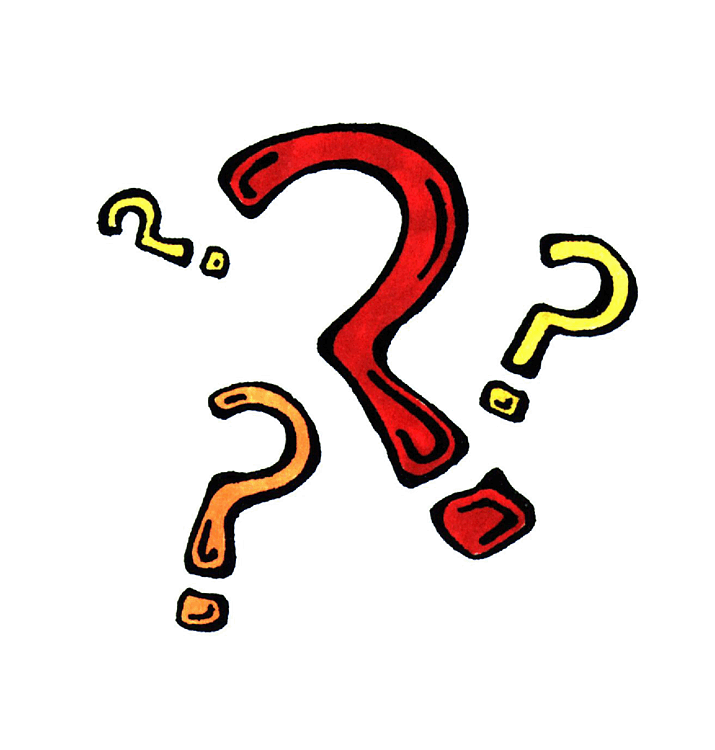
There are a number of responses to this. One simple one is to look at the historical circumstances of Aristotle’s life, where Greek culture as a whole did not divide as strongly between the individual and the community. Life was led much more in public, and as a whole it was much easier to envisage a society where everyone strived to be virtuous against the expectations of others. But we can push this analysis a little further, and argue that Aristotle fundamentally saw human beings as communal animals, and, in his rejection of false ends of moral conduct such as money or pleasure, naturally promoted an ethical system that emphasised reasonability, tolerance, fairness and generosity. In this way, we might not have to state the needs of others as a divide between selfishness and altruism. Instead, for virtue ethics it may simply be a given that the virtuous person is one dedicated towards others, both mentally and physically.

One might also state that virtue ethics is simply being realistic about the ways to encourage people to be moral. If morality was as easy as recognising the need to help others then we arguably would all be saints, but human psychology is more complicated than that. Each person has a sense of self and priorities associated with that self that need to be accommodated in order that a person can arguably feel able to give time and energy to others.

Lastly, we might state that other ethical theories are missing a sense of ambition that is important with virtue ethics. It is simply a given in deontological ethics and utilitarianism that one should follow the rules and principles they set out, whereas in virtue ethics there is the ideal that people should always do more to achieve excellence. This means people can be inspired to develop themselves along with others, and there is not simply the arbitrary, principled requirement to be moral in a particular way based on someone else externally imposing some ethical system.



**Stretch and Challenge**



**Aristotle’s Virtue Ethics: Quick Quiz**

1. Name three of the virtues Aristotle identifies.

2. What word does Aristotle use which is often translated as ‘virtue’?

3. What does ‘Eudaimonia’ mean?

4. The golden mean attempts to find a balance between what two things?

5. What is ‘practical wisdom’ useful for?

Applied Ethics

## Introduction to Applied Ethics – Use of Utilitarianism, Kantian Ethics and Virtue Ethics

Applied ethics as a field is arguably just as important as the pure study of normative ethics itself. For if a moral theory is thought to be successful then it should be able to enlighten what might be the correct perspectives on how to act within any particular moral situation. Furthermore, the very process of using a normative ethical theory to deepen one’s understanding of ethical dilemmas can in turn expose shortcomings in the normative ethical theory itself. For if a particular system of normative ethics is supposed to be exhaustive in informing human action, then it should be able to offer guidance on every moral issue human beings face.

Nevertheless, in practice this rarely occurs. Although changes in ethical thinking have informed progress and understanding of key moral issues, there are still deep arguments as to whether ethics has really informed our moral thinking on certain subjects, or whether we have made ethics conform itself to our already predetermined thoughts. More than in any other part of ethical philosophy, applied ethics has the potential to expose our psychological preferences and biases about a particular issue, and the way our ethical view might not necessarily be based on rigorous thinking about the facts and dilemmas in question.

In this section, four ethical issues will be examined: stealing, simulated killing, eating animals, and lying, with each being analysed from all the major normative ethical positions studied so far. For each you should evaluate which ethical theory provides the greatest insight, and develop your own opinions on how to approach the moral dilemmas raised in each separate issue.

## Applied Ethics 1: Stealing

The question of in what contexts one might be allowed to steal has been a longstanding ethical dilemma that to this day is arguably far from being resolved. On the one hand, we have the idea of possessing objects, such that we have a right to own and do with them as we see fit, yet on the other hand there is the idea of a general right to life and a certain degree of welfare. This means that general dilemmas can be presented where an individual has to steal to survive, and whether such an action is acceptable depending on the laws it breaks or the level of harm it inflicts upon the individual who is being stolen from.

A classic example of a righteous thief is often thought to be Robin Hood, who stole from the illegitimately or unfairly rich to give to those who were poor and could not extract themselves out of poverty. Such a case of stealing is often thought to be justified for numerous reasons, from the immoral actions of the rich in accumulating their wealth, to the social need of those who are poor. Overall, therefore, there is a debate as to whether Robin Hood could be said to be stealing in the conventional sense, or whether he was rightly returning goods to those in poverty. One should be aware from such examples that ethical dilemmas around stealing do not just revolve around the act itself, but also the way we view the rights of people to own objects, the rights of people to live a fulfilling life and the means and ends associated with each of these sets of rights.

#### Utilitarianism

On the face of it the case of whether stealing is right seems to be reasonably easy to resolve from a utilitarian perspective. Simply one can say stealing is acceptable if it produces more happiness than it does unhappiness. Certainly for an act utilitarian this seems to be an acceptable conclusion, and it certainly does allow for cases such as Robin Hood, where instinctively we want to argue that stealing from the excessively rich to give to the poor is allowed. Furthermore, the act utilitarian can claim that in general people will not and should not steal as on the whole it produces more unhappiness than happiness through the erosion of trust in others to respect property rights.

Yet at the same time we earlier analysed how stealing often calls into question the extent to which we own and have rights over our own property. Under act utilitarianism, however, it is very difficult to justify any general human rights or natural rights over and above the total happiness of the greatest number. This means that irrespective of whether I believe I should be able to keep my belongings, so long as someone else will get more pleasure than I will pain by them being stolen, I should not morally object to stealing in principle. This creates a difficult situation, as one could argue that act utilitarianism means that no one really has a right at all to own their belongings, and stealing should be practised whenever it is thought to produce more happiness overall.

#### Rule Utilitarian Solutions

Therefore, one might look to rule utilitarianism in resolving the ethical issues with stealing in order to enforce the general rights of a person to their own property. Yet there are also difficulties with this position, especially when one considers Smart’s criticism that rule utilitarianism reduces down to act utilitarianism.

For example, say we accept a strong version of rule utilitarianism, and establish that stealing is wrong regardless of whether it produces more or less happiness. If this is the case, then we might fall into the trap of ‘rule worship’ again, as we may be ignoring pressing cases such as Robin Hood, where stealing may overwhelmingly produce more happiness than unhappiness. This is especially pressing when one considers unjust economic systems of distribution around the world. In these cases, if regulatory change was not ensuring that a set of ordinary people were receiving adequate payment for their work, and if they were suffering in unjust conditions, they would not be permitted to steal from the excessively wealthy to supplement their meagre income. Rather than help solve Robin Hood cases, it seems as if strong rule utilitarianism simply reintroduces the problem.

However, we can also analyse the issue from a weak rule utilitarian perspective, which would allow for exemptions either when rules clash or when there would be a drastically significant increase in happiness if the rules were broken. This would mean that potential Robin Hood cases would be allowed while preserving property and ownership rights in general cases. While individuals like Smart would contend this is in effect dressed-up act utilitarianism, it can be argued that Mill’s ‘secondary principles’ do provide a good guide on when it might generally be okay to steal.

Overall, therefore, all forms of utilitarianism, apart from perhaps strong rule utilitarianism, provide grounds on which it would be acceptable to steal. However, it also has to be noted that the weight given to property rights, and the happiness associated with them, affects how one perceives the permissibility of stealing. If one believes that, in general, the degradation of property rights from people stealing belongings always brings about more unhappiness than the happiness gained from the theft, as a utilitarian one might be more likely to veer towards a strong rule utilitarian analysis of theft.

#### Kantian Ethics

One might initially think that the approach of Kant to the issue of stealing will be very different to that to the utilitarian, but despite the differences in the way the moral worth of stealing is measured, they are both often concerned deeply with the protection of property rights. In fact, Kant himself argues that stealing is wrong as to conceive of a world where it is permissible would render the possession of any property at all inconceivable. He states:

Whoever steals makes the property of everyone else insecure and therefore deprives himself … of security in any possible property.

Therefore, it is illogical and contradictory to support stealing as a practice. Its lack of universalisability means that there is a perfect duty not to steal in any circumstance, and moreover we can contend that taking someone’s belongings without their permission is treating them and their livelihood as a means to an end, violating the second form.

However, we can once again turn to the Robin Hood example, and here it seems that Kant’s laws potentially prop up unethical economic systems. In this way, it is difficult to know whether Kantian ethics has a real practical application in the world, especially when it comes to complex ethical dilemmas. One might instead formulate the law ‘One ought to steal when it directly benefits the oppressed and poor’, but this arguably still violates certain property rights of those who are rich, a conclusion Kant himself might be wary about drawing. While the moral law is clear and universal for Kant about not stealing, such that all people might have a duty to not unjustly take what is not theirs, it might be contended that not everyone acts according to these rational motives, and a key part of applying normative ethics is understanding how it can determine right action in a variety of scenarios. Therefore, it can be questioned, despite Kant’s clear stance on stealing, whether our moral intuitions about Robin Hood type problems can effectively be resolved.

#### Discussion:

*Would you argue there are situations where it would be acceptable for someone to steal your belongings?   
Does such an argument devalue rights to personal property?*

#### Virtue Ethics

The question of whether it is ever acceptable to steal is arguably more flexible in virtue ethics. We might initially analyse that stealing things is usually done out of self-interest, particularly in relation to vices such as greed. In fact, in *Nicomachean Ethics*, Aristotle argues there are some acts which can never be virtuous, such as murder and theft. Yet modern virtue ethics might argue differently, particular because of Robin Hood cases where we might hold someone as righteous for stealing from the illegitimately rich. In these cases, a virtue ethicist might look towards Aristotle’s emphasis on practical wisdom, and the idea that the right thing or virtue to practise should be judged on a case-by-case basis.

This idea might be illuminated further by considering Robin Hood cases as an example of clashing virtues. While Robin might be acting in a way typically vice-like in stealing from the rich, he is also performing his actions out of a sense of benevolence and perhaps identifying that the only appropriate action is a forced redistribution of wealth to those in need. Furthermore, we can consider cases of partiality as of particular importance in virtue ethics. For example, stealing to feed a member of my own family rather than a stranger is perhaps a more understandable act in virtue ethics, as it exemplifies the virtues of familial compassion, loyalty and friendship. Therefore, it may be the case in modern virtue ethics that stealing is acceptable in some circumstances, so long as it is performed out of virtuous motives rather than in accordance with vice.

Nevertheless, we also have to consider how such acts of theft would relate to Eudaimonia. The difficulties there are in reconciling vice-like behaviour with flourishing have been noted before, but the virtue ethicist might wish to apply limitations to theft, such that the acts of theft do not exemplify virtues in themselves. This might be contained in a full description of practical wisdom. While I might steal food on occasion to feed my relatives, a proper consideration of virtue might also reveal to me that this behaviour is unsustainable, and will eventually mean I pick up a habit of thieving not conducive to my overall flourishing. In this sense, practical wisdom might both reveal the particular circumstances where theft is acceptable, and the circumstances where I need to refrain from it and look for long-term solutions to the ethical dilemmas I might be facing.

#### Discussion:

*Which of the three theories of normative ethics do you think gives the most practical advice when trying   
to decide whether to steal in a particular situation?*

## Applied Ethics 2: Simulated Killing

This is an intriguing topic that has become increasingly relevant in the past couple of decades due to the growing popularity of video games. The key issue that will be discussed here is whether *representations* of killing have any moral significance. This applies as much to television, film and plays as it does to video games, but the latter are a particularly interesting case as so many titles allow individuals en masse to play the role of the killer.

***Discussion:***

*Is there a difference between simulated killing and simulated sexual abuse? What morally significant   
factors could be used to draw such a distinction?*

The relationship between simulated violence and real-life violence is an important issue, but not necessarily for philosophers. Although psychologists seem to have carried out an endless array of studies on the matter, there is still no consensus as to whether exposure to simulated (or virtual) violence gives rise to violence in real life. At best, we may be able to say that exposure to simulated killing makes *some* people more violent *some* of the time. We will see that such findings are more relevant to some ethical approaches than others.



Go to **zzed.uk/9765**

🖰

**Cross-curricular Research Task**

If there are any Psychology or Sociology students in your class, kindly ask them to briefly outline what   
they know about media violence research. If not, visit this link (*http://www.simplypsychology.org/bobo-doll.html*) to learn about one significant study in the field.

Nevertheless, the assumed ability for graphic depictions of violence to have a corrupting influence on society is a perennial cause ofmoral panic among the British tabloid press, as evidenced by the media debate that accompanies each new release in the Grand Theft Auto franchise. This is a trend with ancient lineage; the idea of art more generally being somehow ruinous can arguably be traced back to Plato’s *Republic*, where it is held that music, poetry, drama and painting are dangerous due to their illusory nature and tendency to stir the emotions. This brings us to a different point: is the experience of simulated killing *bad for us*? Conversely, could experiencing simulated killing actually be *morally good*? This is a question that belongs as much to aesthetics as it does ethics, and so it brings with it the opportunity to survey some philosophical perspectives on the arts.

Our chief areas of interest, then, will be the following:

• Is simulated killing ethically significant?

• If it is, is observing or participating in simulated killing necessarily morally *bad*?

#### The Utilitarian Perspective

What is wrong for the utilitarian is that which, on balance, causes more suffering than it does pleasure. Does simulated killing cause suffering? There are two senses in which this could be true; the first, which has been alluded to already, is the possibility that simulated killing could lead to real killing. Resolving this question is likely to be of greatest importance to the utilitarian as they are concerned solely with *consequences.* Yet, empirical studies into the relationship between violence in video games and real-world violence have failed to establish a conclusive link. Even if a link were proven, the utilitarian would still have to weigh up whether the suffering caused by those real-life killings outweighed the pleasure created by the simulated killings.

A second sense in which simulated killing could cause harm is by causing emotional distress to the viewer or player of the simulation. So-called ‘griefing’ in online video games, where players purposely try to ruin the experience of other players by, for example, killing their teammates, could be an example of this. Similarly, the simulated killing of popular characters in TV shows and film can also cause fans serious distress (see, for instance, the hysterical reactions of some viewers to deaths in *Game of Thrones* or *Harry Potter*). As with the first sense, the utilitarian must still decide whether the suffering caused by the simulated killing outweighs the pleasure (and this is unlikely to be true).

It is worth at this point reiterating Mill’s harm principle from *On Liberty*: *‘The only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others.’*

Practically speaking, if simulated killing were proven to cause harm to other members of the community then the only solution, for the utilitarian, would be to restrict access to certain films, games, TV shows and so on. Yet if this cannot be established, and we have good reason to think it cannot be, then the utilitarian is likely to reject the argument that simulated killing is morally wrong; Mill, in particular, would baulk at the idea of censoring free speech.

**Aesthetics**

The branch of philosophy which deals with questions such as: are certain works of art objectively good? Does art serve a purpose? How do we form judgements about what is and is not an   
artwork?



An alternative consideration is whether simulated killing may actually offer benefits in terms of utility; for instance, stress relief or plain old fun. From Mill’s perspective, the value of simulated killing would likely be tied to the value of the pleasure. Although Mill did not have access to computer games or films, he would certainly have had access to works of art which contain simulated killing. Being able to appreciate the (numerous) acts of murder in Shakespeare’s works would likely constitute a higher pleasure for Mill. It seems, though, that very quickly we begin to enter the realm of **aesthetics** and something of a troublesome area for Mill, who has more than once been accused of being elitist. The accusation is hardly surprising considering Mill’s social and historical context (essentially an upper-class Victorian gentleman). It is to be expected that Mill would have considered watching Michael Bay films or playing Grand Theft Auto as a far inferior pleasure to a nuanced understanding of the works of Goethe or Byron.



Nevertheless, we must not disregard the higher–lower pleasure distinction out of hand simply because we live in an age where taste is largely considered a subjective matter – ‘different strokes for different folks’ as they say. If we can say that, for example, people who have watched a (notoriously violent) Quentin Tarantino film would prefer to watch another of his films rather than slog through another third-rate, straight-to-DVD slasher flick, then we might have grounds to say that experiencing simulated killing can actually be considered a higher pleasure and so is *morally good*.

#### Pleasure vs Preference Utilitarianism

One distinction we can potentially make in arguments about simulated killing is that between pleasure and preference utilitarianism. In the former, we’ve noted the difficulties in trying to understand whether simulated killing might cause overall harm, but preference utilitarianism arguably simply resolves the issue much more easily. For if we generally decide we have a preference that corresponds to a desire to play violent video games, then even if they potentially have an overall negative effect on our mental and physical health, we should be allowed to play them. This is particularly interesting in the context of Mill’s liberty principle, for it might be posited that violent video games would be something that Mill would not want to outlaw for fear of violating overall liberty. But the fears over the harm they can cause may well prevent their continued legality, and it can be argued that the liberty principle as a whole may be better supported under a preference utilitarianism framework.

***Discussion:***

*Do you think differences of taste undermine the higher–lower pleasure distinction? Can you think of an   
example of simulated killing which could be considered a higher pleasure? What is wrong with the argument I just presented?*

#### The Kantian Perspective

It is difficult to see how fictional characters in films or TV shows could have any moral status for a Kantian. As with animals, their lack of reason means they are ethically insignificant. We might speculate, however, about simulated killing in online video games. Does killing another player’s avatar in-game involve treating them as a means to an end? After all, the person behind the simulated killing probably enjoyed giving someone a simulated blast in their simulated face with a simulated shotgun, and the player controlling the simulated victim was perhaps a little miffed. In this sense, one player is being used as a means to the other’s pleasure. Yet this is no different to any other competition; there will always be winners and losers. Kant explicitly states that we have an imperfect duty to develop our ‘natural talents’. The maxim being acted on in a video game may be something like ‘use your natural talents to be the best player in the game’ which, when universalised, involves no contradiction.

***Discussion:***

*Does competition always involve treating others as a means to an end?*

The Kantian may, however, wonder whether overexposure to simulated killing damages our capacity to exercise the good will. Remember, what is wrong for Kant is what is *irrational*. Murder is wrong not because it causes pain and suffering, but because it violates the formulations of the categorical imperative – ‘Kill ’em all’ is not a universalisable maxim, nor does it treat others as ends in themselves. Is, therefore, wanting to experience simulated killing an irrational desire?

However, just as Kant argues that human beings should generally treat animals well so as not to develop a bad habit in treating human beings badly, it can be argued that people should not play violent video games if it is the case that this may lead to them harming human beings more easily in real life. Part of acting out of duty involves developing oneself such that one is not tempted to act out of self-interest, and it can be argued that the simulation of violence may lead to a person actually committing violent acts in real life.

Kant argues as such that we have an imperfect duty not to participate in or observe – for pleasure – simulated killing; perhaps no rational being would truly want to derive satisfaction from an immoral action. This may arise in particular out of a perfect duty to perhaps cultivate compassion for living beings, or developing the natural ability to treat other people as ends. On the other hand, if we observe or participate in simulated killing because it illustrates some higher moral purpose, the triumph of good over evil say, we do not face the same problem. As ever for the deontologist, it is the intention behind the action which really counts.

#### The Aristotelian Perspective

Of utmost significance to the virtue ethicists is the extent to which an individual is achieving Eudaimonia. So, the Aristotelian may wonder whether a person who spends every waking moment slaughtering digital Nazis on Call of Duty or amassing virtual currency by slaying dragons in a world of pure fantasy is really an example of someone leading a flourishing life. The same point can be applied to those obsessed with ultra-violent films or television shows. This is likely a question of moderation; too much time devoted to what one finds enjoyable may lead one to ruin. In addition, being wholly desensitised to violence could lead to a hardness of heart which might make one more prone to maliciousness or insensibility.



An alternative consideration is whether the experience of simulated killing could somehow contribute to Eudaimonia; in particular, it may be *cathartic*. While Aristotle did not have access to films, television shows or video games, he did have access to the theatre. He was especially interested in a genre of Ancient Greek theatre known as tragedy, and devoted much time to it in a work known as the *Poetics*. The word ‘catharsis’ itself takes on much of its modern meaning from Aristotle’s discussion; it is something which somehow purges or purifies us of certain emotions. In the case of tragedy, where the audience views a simulation of great suffering or misfortune, the feelings of fear and pity which overcome us during the performance are said to be purged by the end. This process gives rise to a peculiar pleasure, a sense of being purified or even of having greater insight into the indefinable ‘human condition’ (IEP, *Poetics*).

We might apply this notion to simulated killing. Here the violence acts as a purging mechanism, bringing up the urges that lurk in the pit of our stomach and letting them out into the open. In this way, experiencing simulated killing may rid us of what Freud called *thanatos* – the ‘death drive’, which urges us towards aggression. In this way, an Aristotelian might claim that certain kinds of simulated killing actually lend themselves to Eudaimonia, or at least move one away from the vices of irascibility and cantankerousness.

However, this is perhaps a difficult view to justify. In his paper ‘Is it wrong to play violent video games?’, the philosopher Matt McCormick actually argues that virtue ethics is the ethical system that perhaps provides the strongest opposition to simulated violence, for it can be strongly argued that acts such as virtual murder erode the possibility of cultivating virtue in the wider sphere of one’s life and achieving Eudaimonia. This naturally relies on the link between cultivating a violent disposition in-game, and cultivating a violent disposition outside of it, but the link might also be more subtle. All it would take is to say that the general inwardness and isolation that may accompany video games is not conducive to achieving flourishing. If I spend all day running down pedestrians in Grand Theft Auto, I can hardly be said to be developing long-term habits to improve my own character. Therefore, virtue ethics might provide the strongest opposition to simulated killing as a general practice.

## Applied Ethics 3: Eating Animals

The question of whether it is acceptable to eat animals is one that has been magnified in the last 100 years. While various societies have had significant dietary restrictions, the ability in the modern era to have a balanced diet without the consumption of meat means that many people, especially philosophers, have questioned whether it is right to eat meat if not necessary. Such questions have also been bolstered by new scientific evidence that animals have a much richer inner life than previously supposed, being able to think, feel, and experience pain in greater ways than was thought by most individuals throughout history. This raises distinct moral issues, especially when one considers that much of the meat in the world is now raised in factory farm conditions, which often have poor standards of hygiene and force animals to live in conditions far from their normal experiences in the wild. For example, in the case of chickens, modern farms often include in their feed a significant amount of antibiotics that causes them to mature and be ready for slaughter in just 40 days, whereas chickens in the wild might expect to live (if surviving childhood) for an average of seven years. Furthermore, male chickens are slaughtered at birth, as they are not suitable for laying eggs or for consumption. Such practices not only give rise to questions about what sort of standards animals should have if farmed, but to what extent animals deserve to be farmed and eaten at all.

Another issue has also been to what extent our consumption of meat has a deep environmental impact on the Earth. Raising animals by and large takes a much greater amount of land, energy and water than growing crops, and it can be argued from a case of self-interest that human beings need to cut down on their meat intake to ensure that they continue to thrive and prosper into the future.

Therefore, there are many distinct issues raising by eating animals, and different ethical positions may offer insight into whether human beings should cut down on eating animals, or even stop eating them altogether.

#### Discussion:

*What do you think are the most important ethical concerns in eating animals? Is it the rights or welfare   
of the animal itself, or the overall impact meat eating has on human life and the environment?*

### The Utilitarian Perspective

#### Classic Utilitarianism

The classic pleasure/pain-focused utilitarianism may look at meat eating from a number of angles. The first is simply to say that only human happiness really matters, and unless eating animals causes human beings severe distress, then it is likely that eating meat increases overall happiness and is, therefore, acceptable. Certainly this is the case if one is considering eating animals where there is a general scarcity of food. On the whole, for a utilitarian, the welfare of human beings may be prioritised especially where survival is at stake, as this affects a great deal of potential future happiness that might accompany a living human being.

#### Singer’s Utilitarianism

However, it was noted in the utilitarianism section that other forms of utilitarianism exist. Singer, for example, argues that moral weight should be given to the pleasure and pain of animals, such that only focusing on the happiness of human beings is being guilty of speciesism. Note this does not mean that Singer argues that we should not eat meat. In fact, even though he is a preference utilitarian, he argues that on the whole animals do not have the rational capabilities in order to have preferences that we can understand and identify. Therefore, we are not obliged to adhere to animals’ potential preferences in the same way that we do human beings. However, a case can still be made that if human beings do not need to eat animals to survive, then there is a pressing case to reduce overall suffering by at least reducing our general consumption of meat, if not stopping it altogether. Therefore, there is a fine balance in not over-prioritising the concerns of animals, but at the same time maximising utility by minimising the suffering animals have to undergo in order that we eat effectively and survive.

#### Difficulty of Assessing Overall Impact of Meat Eating

Two different forms of utilitarianism have been looked at so far in the discussion around eating animals, yet it can be seen that overall there is a difficulty to analysing the overall impact of and suffering caused by consuming meat in general. In fact, you might remember the issues surrounding calculation of suffering in utilitarianism from earlier sections, and eating animals is a classic example of how utilitarianism, despite being simple at its core, might struggle to deal with complex ethical issues. For example, should we weigh the suffering of chickens differently from that of more intelligent pigs? If so, then there could be an argument that we should eat more chickens instead of pigs. Or should we prioritise our future suffering from the environmental damage of factory farming over our current pleasure from enjoying its spoils? Utilitarianism may provide the basic tools to assess this problem, but whether the qualitative pleasures and pains of eating meat can be compared is another question entirely. It may be the case that two utilitarians vastly differ in their conclusions on the ethical viability of eating animals.

#### Act or Rule Utilitarianism?

One thing that can be asserted, however, is that, in the case of eating animals, it seems that act utilitarianism simply cannot accommodate the wide variety of factors and influences that govern meat eating. For example, how do I know that the steak I am eating comes from an animal who has had a happy life, did not meet excess suffering at its end, and did not contribute excessively to the environmental damage factory farming causes? It seems at a glance, if one is not aware of such issues, then it is impossible to really judge whether any particular act of meat eating is correct, except perhaps when pure survival is concerned. Therefore, there may be a strong argument that some sort of utilitarian rule is necessary when discussing meat eating, and that human beings do not have the conceptual tools to discuss meat eating on an act-by-act basis.

### The Kantian Perspective

#### Second Formulation, The Primacy of Reason

Kant stresses the important role of reason in morality throughout his works, and, as has been explored in the section on Kant, reason grounds the ability to be moral in the autonomy of human beings. But Kant argues that animals do not have freedom or reason, and so cannot treat others as ends. They can only will what they desire through instinct. In this way, for Kant we are not required to treat animals as ends as they lack the fundamental rational capacities to engage in the process of making moral laws and following them.

What does this mean? Well, in short, human beings are allowed to eat animals under Kantian ethics. Animals do not have the moral status deserving of being treated as ends so are allowed to be treated as means to either our survival or pleasure. However, Kant does stress that human beings on the whole should seek to treat animals as well as possible, as, if we get into the habit of treating animals badly, we may transfer that habit towards treating other human beings badly. Furthermore, although animals do not have the same status as human beings, they may still have moral import if we universalise laws that require us to be compassionate to all beings, or not create excess suffering. Therefore, although it is strictly acceptable to eat animals for Kant, there may be wider moral principles governing to what extent human beings can indulge in meat eating (in fact, Kant himself envisioned as part of their duty towards developing moral self-betterment, human beings have an imperfect duty to not treat animals cruelly).

#### Issue with Human Beings?

One issue you might pick up on here is that Kant’s argument effectively means that any being without reason can be treated as a means to an end. This could include babies or those who are mentally disabled, and many critics have not taken favourably to this element of Kant’s ethics, for we often suppose that we have distinct duties to those who are unable to take care of themselves, or are incapacitated in some way. Furthermore, it might imply that if we ourselves become temporarily mentally disabled – for example, in an epileptic fit – we are allowed to be treated as a means until we recover.

This is a difficult issue for Kant to resolve so long as he attaches moral worth purely to the ability to be rational, but he may reply similarly to the issues surrounding animals; that human beings on the whole have a duty to not be cruel or cause excess suffering. In this way, we might recognise that although those who are mentally disabled do not have the same strict moral status as rationally capable human beings, we still have moral duties to them based upon the requirement for each person to dedicate themselves to their own moral self-betterment.

#### Animal Intelligence

Another interesting issue is whether we can revise Kant’s conclusion with our growing scientific knowledge about the intelligence and rational capabilities of animals. For example, chimpanzees have been taught hundreds of words, can arrange them in grammatically meaningful sentences and can communicate with people about their feelings, desires and interests. Does this mean they effectively qualify as rational beings under Kant’s second form? If so it may mean that we have distinct moral duties not to eat animals that display a certain level of intelligence, which may include everything from pigs to octopuses and dolphins.

### The Aristotelian Perspective

#### Discussion:

*Do you believe there is a more of a moral imperative to avoid eating animals that have higher levels of intelligence and consciousness? Or should we simply avoid eating all animals?*

#### What does a virtue ethicist need to consider?

Compared to the utilitarian and deontological approaches to the ethical issues around eating animals, virtue ethics arguably approaches them from a different perspective. In both previous sections, we’ve explored how the moral status of animals, often based around certain characteristics they do or do not possess, influences the moral regard we are supposed to give them. While Aristotle himself was guilty of ordering living things into a hierarchy, with the reason-possessing humans at the top, virtue ethics as a whole is arguably more concerned with how eating animals might reflect on an overall virtuous character. Therefore, for a virtue ethicist, eating meat in any situation needs to be considered against not only the potential suffering of the animal itself, but also against the overall virtues a person should possess. So the question is opened up about whether a person who could be described as compassionate, benevolent or righteous should inflict excess suffering upon other living beings.

#### Hursthouse and Vegetarianism

You might remember how Hursthouse criticised Singer for his one-sided focus on the rational and conscious capabilities of animals, ignoring the wide variety of roles they play in human life, and the various categories of language human beings use to describe them. In fact, she argues that focusing on the moral status of animals is something that virtue ethics does not need to do, as it approaches the question of eating meat from a more holistic perspective, analysing not just how meat eating affects animals, but also our character and ideals. In fact, only focusing on the moral status of animals is a misleading approach, as there is no specific set of characteristics that separates animals in general from humans and even other animals. Trying to order our meat eating preferences by a hierarchical scale of sentience is anything but objective, and simply introduces a misleading way of categorising the moral worth of our actions by assigning moral worth to something separate to what should be the focuses of our ethical decisions: virtues and the duties and responsibilities that come with them.

#### Virtues and Eating Meat

So, if we are considering eating meat, what virtues might this action exemplify? Well, many philosophers have argued none, and, in fact, the act of eating meat to a great extent is an uncompassionate act for it involves cruelty in inflicting suffering upon an animal. For example, if we consider factory farming, many animals are kept in cramped, unhygienic conditions and often force-fed to induce quicker growth. It can hardly be said that encouraging these sorts of practices is exemplary of the virtue of being compassionate. Therefore, there is generally a clear argument against eating meat raised by unethical farming practices.

But what about meat gathered from hunting animals in the wild? Or raised on free-range, organic farms? This is a trickier question to answer, for it could be argued that on a farm dedicated to high standards of welfare for the animals, there is great compassion in ensuring the animals have a healthy and enjoyable life. One might argue that one is committing no extensive vices by eating animals raised under such conditions. Yet Hursthouse argues that eating meat from these farms is still eventually committing animals to suffering and death by human hands, and so it can hardly be said there is compassion in these acts, especially when human beings do not need to eat meat to have a healthy diet. In this sense, if any unnecessary suffering is caused to animals simply so that we can consume them, we can hardly be said to be exemplifying compassion and virtues within ourselves, making eating meat categorically non-virtuous.

But one more thing can be said in regard to meat eating and virtue ethics. If you recall, practical wisdom is the most important process for a virtue ethicist in deciding what moral action to take, and it can be hypothesised that there may be occasions where a human being is faced with eating meat or starving. Here it still may be justified under virtue ethics to eat meat, even if it is not a virtuous act itself. As Aristotle states often, becoming virtuous is a difficult process, and is dependent on material conditions as it is a person’s own engagement with habituating virtuous behaviour in themselves. In this way, virtue ethics does not proscribe distinct rules about not eating meat, but analyses it on a case-by-case basis. Those with the means and will to refrain from eating meat will probably see it a virtue to be a vegetarian or vegan, while those who need to eat meat to survive will do so through considerations of practical wisdom.

## Applied Ethics 4: Lying

Whether it is acceptable to lie is an age-old question, and has been discussed in ethics perhaps more than any other topic. It covers numerous areas, from whether the consequences of our actions matter more than the intentions, and to what extent there is a general value on truth. What is greatly difficult about discussing whether lying is ever good is that the nature of a lie depends greatly both on the context of the situation and the intentions of the person involved. Nevertheless, looking at the ethics of lying from the normative ethical positions so far can reveal interesting aspects about this issue, and how we may consider cases of lying in our own lives.

#### The Utilitarian Perspective

A pure **act utilitarian** will decide that a lie is the correct course of action if it produces the greatest amount of happiness for the greatest number of people. Conversely, if a lie is likely to produce more pain than it will pleasure it is the wrong course of action.

We have already touched on some of the difficulties with this kind of thinking. Firstly, a lie may have unintended consequences. Consider the old morality tale: I tell a small fib today to get myself out of a sticky situation, but the next day I must tell a slightly bigger fib so that nobody discovers my first lie. The next day the same thing happens again: a bigger fib still to cover up for my first two lies. It goes on like this until I have spun such an enormous web of deceit that I wish I had simply been honest in the first place. The moral of the story? Just tell the truth.

A similar argument can be used against the act utilitarians’ justification for lying. You may calculate that it is the action likely to produce the greatest happiness, but you cannot know what effects it will have in the weeks and months to come (this is a problem for consequentialist theories generally).

Secondly, if deception can be justified in this way, it may permit some fairly objectionable courses of action. Here is one scenario: a depraved murderer is on the loose in a small town and the residents are growing increasingly panicked, if not hysterical. Local papers run with overblown headlines such as ‘Police are Powerless to Stop Sick Killer’ and fretful parents are pulling their children out of school. A pure act utilitarian in this situation may recommend imprisoning a harmless oddball for the murders despite knowing full well that they had nothing to with the killings. Knowing that the police had got their man and the streets were safe to walk again at night would no doubt bring great pleasure to the terrified townsfolk, pleasure enough to far outweigh the pain felt by the harmless oddball who is being punished for a crime he did not commit.

Yet while a utilitarian might initially support such a conclusion, it can be argued that it would be hard to reach, for inbuilt in lying are a number of negative conditions. The more people lie, for example, the more trust is eroded, and so people on the whole become more unhappy when they can’t live their lives without worrying that they’ve been lied to. In the example of the police, they have to consider that the murderer is still on the loose and may kill again, and that it may be found out that the person they imprisoned is innocent, making the situation worse than it is before.

However, we can also state that for a **rule utilitarian** this problem does not apply. As we saw earlier, rules which preserve certain rights (such as ‘The right to not be falsely imprisoned’) will generate greater happiness than allowing people’s rights to be ignored when it is convenient.

It is less clear-cut how a rule utilitarian would deal with the first problem.

On the one hand, they could adopt a rule such as ‘Do not lie’ or ‘Do not deceive’. This would avoid the unintended consequences of telling lies and deceiving others.

However, it would also mean that the rule utilitarian would fall victim to the problems which plague Kant’s theory (discussed below); chiefly, it would mean that a lie would be impermissible even if it were *clearly* the right course of action. The case of the inquiring murderer is a classic example: if telling the truth is likely to result in the death of an innocent, then, all things considered, lying is the best course of action.

What the rule utilitarian might say is that some safeguarding principle could be created which means exceptions to the other rules can be made in particularly grave cases. However, this then leads to the question of how we judge a case to be grave enough to warrant an exception. These difficulties may not be unsolvable, but they can lead to such levels of complexity that the entire approach may begin to look wrong-headed.

We can, however, despite these issues state the advantages that utilitarianism brings to telling lies. For example, parents will often lie to their children that Santa Claus exists, for it brings the children more happiness, and the unhappiness they experience when they find out he is make-believe is much less on the whole as they have matured to the point where they can understand why the lie was told. In this way, while we have examined the ways in which long-term lies can prove difficult for utilitarianism, conversely the ethical theory also allows us to realise why a short-term lie might be good, even if exposed later.

#### Preference Utilitarianism

We can illuminate this idea further by considering preference utilitarianism and lying. While it is generally obvious throughout my life that I should have a preference for people to tell me the truth, I can also hypothesise that there are times when I would rather have people lie to me. For example, if I write my girlfriend a song for her birthday, no matter how bad it is, I would generally prefer to hear good feedback, especially if it was not intended for anyone but her. In this situation, I may well say I have a preference to be lied to rather than told the truth. Therefore, while the issue of lying is complex, it can be argued that utilitarianism gives us the tools to judge ourselves when a lie might be acceptable, not just in the context of our lives, but others as well.

#### Discussion:

*One issue which remains in utilitarian theories is that of calculation. Can we ever be sure that a lie won’t have severe unintended consequences, as is this a reason to believe one should never lie?*

#### The Kantian Perspective

We have already discussed the problems the telling of lies can cause a Kantian when we discussed the case of the inquiring murderer.

Kant’s position is fairly clear on this; we have an absolute duty to tell the truth and not to lie. Following the first formulation of the categorical imperative, any maxim which permitted lying could not be made into a universal law and so could not be considered a good act. Following the second formulation of the categorical imperative (the humanity formula), deceiving or lying to somebody is to treat them as a means rather than as an end. It is essentially an affront to their rational nature. In Kant’s view, people should be respected enough to be told the truth so that they can make a fully informed, rational decision.

Yet from the issues presented by the case of the inquiring murderer we can also ask whether a duty could ever be formulated where lying is acceptable; for example ‘one should lie to protect the life of another’. However, as noted before, it is hard under Kant’s system to justify any lies, as they will always contradict the second form. For lying effectively prevents anyone from making rational choices, since they will be at some point accessing untrue information in their deliberations. The upside is that our personal integrity is never violated, but the downside is that we are left with a wholly inflexible ethical system that never lets us act in what we might perceive to be the better interests of the individual in question.

In total, although Kant envisions the first form and the second form as the same in principle, the first form allows us to potentially develop duties for when it is acceptable to lie, while the second does not. So, depending on how one develops Kantian ethics, one is perhaps obliged to choose one over the other, unless one argues that one can lie to an individual without treating them as a means. This may be true, for example, if one knew that a friend would in a particular situation wish to be lied to, meaning that by lying one is fulfilling their rational interests, not just your own. However, such a compromise is still largely inflexible, as we can hardly know for sure the rational interests of our friends, let alone strangers.

Nevertheless, one way to envision why Kant has such a strict opposition to lying is in reference to his additional formulation: ‘*Act according to maxims of a universally legislating member of a merely possible kingdom of ends.’* This means that when we consider how to act we should not just think about ourselves, but how our actions would be in a hypothetically perfect moral society. In this way when we lie we are perhaps not just immediately affecting our surroundings but denying the very possibility of a society where everyone treats each other as ends. One other way to think about this is that if in any situation you do not take the moral high ground, then a society where everyone takes the moral high ground will never be possible. Either everyone has to adhere to duty, or the very trust in people to act out duty collapses. Therefore, we have to consider not just ourselves, but the universal implications in lying to others, and the very effect lying has in moving human beings further away from a kingdom of ends.

***Government House Utilitarianism***

One of the more controversial ideas that has been put forward by utilitarians is the so-called 'Government   
House' theory of the Victorian philosopher Henry Sidgwick. According to Sidgwick, a society where   
every person lives by the tenets of utilitarian morality may be neither possible nor desirable. Instead,   
Sidgwick proposes a two-tier system of morality where policymakers and other influential members of a society base their decisions purely on the principle of utility, while the vast majority of people live by the same customary ethics that they always have. This is because the 'vulgar' masses may be unable to understand the principle and so apply it incorrectly. Thus, greater happiness is more likely to be achieved by keeping most of the population in the dark about utilitarianism. So he is also advocating large-scale deception on the part of the government, who must, to maintain social order, lie to the populace.

Unsurprisingly, critics were quick to point out the elitism inherent in Sidgwick's position. By reserving the utilitarian ethic for the 'enlightened few', Sidgwick reveals the prejudices of the colonial age – a time when it was thought the 'civilised' European powers knew what was best for their 'uncivilised' subjects. Nevertheless, Sidgwick's position is more than a historical curiosity.



**Stretch and Challenge**

#### The Aristotelian Perspective

Aristotle only seems to mention virtues and vices which involve *self-deception* in the *Nicomachean Ethics,* namely the vices of ‘boastfulness’ and ‘mock-modesty’ and the virtue of ‘truthfulness’ (in self-expression). The idea here is that it is a mistake to overemphasise one’s abilities and also to understate them. The mean between arrogance and a lack of self-belief that the virtuous person embodies is truthfulness, the ability to give an honest account of one’s strengths and one’s weaknesses.

When it comes to deceiving others, it seems that Aristotle would probably take a much more flexible approach than Kant. On some occasions, the most virtuous act may be to lie (for example, a white lie to protect a person’s feelings); on others, failing to tell the truth may be considered a vice. It can also be argued that certain deceitful actions (such as adultery or theft) arise from vices in a person’s character; for instance, perhaps they are envious of what others have, or they may be self-indulgent. Ultimately the rightness or wrongness of any particular act of deception will have to be submitted to the usual criteria: the virtuous action will be done ‘at the right time, toward the right objects, toward the right people, for the right reason, and in the right manner’ (1106b21–23).

However, in this idea we can also explore how virtue ethics potentially accommodates the different ways in which we lie, not just the lie itself. The classic example that can be given is that of the ‘white lie’, a small or unimportant lie we tell to ease life, or spare someone else’s feelings or stress in an everyday situation. This often involves just withholding information rather than actively deceiving someone. For example, if a friend had a bad day, and I was supposed to deliver bad news about one of their exams they had failed, I might wait until they were feeling better before telling them. In this way honesty is a gradient, and there might be important ideas such as tact and care involved also in telling the truth.

If you remember, therefore, from the table of virtues, truthfulness was listed instead of honesty. In many ways this can represent the need to cultivate honesty, but also the need to recognise what form honesty is best delivered in as per the demands of practical reason. A lie in some circumstances may be acceptable if one is habitually honest, while in other cases it may be the result of laziness or self-interest, and the best time to tell the truth is as much a case of practical wisdom and contemplation as it is simply being honest whenever possible.

Meta-ethics

## Introduction to Meta-ethics

So far in this guide we have studied questions of what people should do in any given situation. However, throughout we have also looked at how different philosophers arrive at their conception of the good, and why they think that they can equate what is good with certain ideas such as pleasure or virtue. These sorts of arguments do not fall under the purview of normative ethics, and are best described as meta-ethical questions. For, rather than looking at what is right to do, they question the very concepts and words we commonly use within ethics, and try to define what they actually mean and how they should be used in philosophical argument.

Therefore, beyond questions about what we ought to do, there are fundamental questions about what moral judgements or statements even are. Such is the purview of meta-ethics, and there are wide-ranging discussions about the status of ethics itself; whether we can definitively say there are moral facts, or whether somewhere along the way we’ve made one assumption too many, and the entire foundations of ethics itself are unjustified. Such thinking has even led some philosophers to deny that moral statements have any objectivity altogether. Rather they are only relative to a particular viewpoint or culture, or even are fundamentally wrong as statements no matter their content.

So how do philosophers go about trying to clarify such issues? Well, one starting point has often been to look at the linguistics and semantics of moral statements, trying to find out to what extent they differ from normal descriptive statements about the world, and whether these differences indicate they should be held in a different regard. Another starting point has also been to question the potential origins of morality, and whether it is a human construct designed to bring order to societies. This has led to some arguing that morality may be boiled down to self-interest, or at least that it is a way of negotiating different interests between competing parties. Others have argued that morality arises of our different cultures or traditions, and considerations about what they consider to be strengths and weaknesses in human beings.

However, whatever the starting point, it is important to be careful when studying meta-ethics. Often we feel that we can explain something away by stating its origins, but there are many things in the world that are still meaningful and true, even if their origins are less than reputable. For example, I might have a friend who refused to study for a multiple-choice exam, wrote down a random number for each answer and still managed to get full marks. Even though they have no idea about how they achieved their result, they still had a correct result. In the same way, although the origins and meaning of our ethical terms may be fuzzy, and it might be possible to give explanations about how they arose, this does not necessarily indicate that they are wrong altogether. There is disagreement in all fields of study, and the mere presence of disagreement does not make a particular idea false.

Nevertheless, throughout this section we will look at a number of key questions:

1. Do moral statements make claims about an external or mind-independent reality?

2. Are there any objective moral facts?

3. What do we mean when we use terms such as ‘good’ and ‘bad’?

4. Can terms such as ‘good’ and ‘bad’ be reducible to natural properties?

5. From where do our moral principles originate?

Ethical Language 1: Cognitivism

The study of ethical language, often referred to as meta-ethics (which means *after* ethics), is the branch of philosophy which studies the meaning of words with moral significance, such as good, bad, right, and wrong.

We will be studying two contrasting approaches:

**• Cognitivism**, which holds that ethical statements make claims about *mind-independent* *reality* and so can be true or false.

**• Non-cognitivism**, which holds that ethical statements do *not* make claims about mind-independent reality and so *cannot* be true or false.

The field gained popularity after the so-called ‘linguistic turn’ in philosophy at the start of the twentieth century. It was particularly popular at the universities of Oxford and Cambridge; in fact, most of the thinkers we will be studying held academic posts there at some point in their lives.

Ethical cognitivism refers to the belief that moral statements are ‘truth-evaluable’; they make claims about the world that can either be true or false.

Closely related to the notion of cognitivism is the position of **moral realism**, which holds that ethical statements, if used accurately, are an objective description of facts.

Moral realism is often associated with **ethical naturalism**, a position in ethics which holds that ‘good’ can be defined in reference to things in the natural, empirical world. Good, for example, can be defined as ‘happiness’ (as in the case of utilitarianism). The statement ‘honesty is good because it brings about happiness’ is a statement that refers to happiness, which is a recognisable thing in the world, and we can check, through empirical evidence, whether happiness was indeed brought about by it. For a utilitarian, if honesty is seen to bring about happiness in the world, then the statement ‘honesty is good’ is not only a moral statement, it is a **fact-stating** statement too.

Moral realism, therefore, holds that there can be factual knowledge about morality.

Not all moral realist theories are naturalist theories. **Intuitionism**, outlined below, is a prominent example. It is a cognitivist theory but it is non-naturalist (does not adhere to ethical naturalism) because, although it considers ‘good’ to objectively exist, it cannot be defined in reference to the natural world. Nevertheless, it is a *fact* that certain things are good.

There is also a peculiar school of cognitivist thought, known as **error theory**, which maintains that, although ethical language does make claims about reality, it so happens that those claims are *always* false. Hence, it rejects moral realism.

That is a lot of terminology to take in, but it will become clearer. This diagram might help:

**Cognitivism**

*Ethical language makes fact-stating claims about reality*

**Ethical Naturalism**

*Ethical language makes true claims features of the natural world.*

**Ethical Non-naturalism**

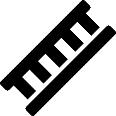
*Ethical language makes true claims about things that are not part of the natural world.*

**Error Theory**

*Ethical language only ever makes false claims about reality.*

**Moral Realism**

*Ethical language makes true claims about reality.*





***Direction of fit***

One way of understanding the distinction between cognitivism and non-cognitivism is to talk about the   
*direction of fit* between the mind and the world.

**Cognitivism** has a *mind-to-world* direction of fit. Cognitivists think ethical language expresses beliefs about how the world is. If I *believe* (in my *mind*) that murder is wrong, it is the *world* which makes it true or false. Think of a jigsaw puzzle that’s missing a piece:

If the piece is correct, it will fit into the jigsaw.

If it is wrong, then we must change the piece if we want it to fit into the jigsaw.

That is how cognitivism works. If a belief is true then it will ‘fit’ into the world.

If it is wrong, we must change our beliefs.

That’s why cognitivism is an *objective* theory.

Your belief, your piece of the jigsaw, is either right or it’s wrong.

It doesn’t matter whether you *personally* think you’re correct.

If the piece doesn’t fit, then it’s the wrong one.

**Non-cognitivism** has a *world-to-mind ­*direction of fit. Non-cognitivists think ethical language expresses how we wish the world would be.

The *world* contains murderers, and I *wish* (in my *mind*) that it did not.

Think about the jigsaw again:

If the piece does not fit, it is not the piece which is wrong but the jigsaw.

The piece tells others how we wish the jigsaw would look.

There is no objectively right or wrong way the world should be.

We each carry around a piece of our ideal jigsaw.



**Stretch and Challenge**

Moral Realism

As we have noted so far, moral realism asserts that there are objective moral facts. This might seem like a bold claim, for if there were moral facts surely we should have access to them? Surely there wouldn’t be so much disagreement about the nature of morality if there were facts?

Yet, as was discussed in the introduction to meta-ethics, disagreement doesn’t necessarily mean there isn’t a truth underneath. There are plenty of people who still argue that the Earth is flat, yet we don’t say simply because there is disagreement that there is not a definitive fact about the shape of the Earth, and, funnily enough, if we looked at the Earth at near the speed of light, then due to the contraction of objects that occurs at those speeds, the world would indeed look flat from our perspective.

This is of course a silly example, but it underlies a common misconception about what moral facts might mean. The fuzzy nature of ethical discussion means that the idea of moral facts often conjures up images similar to Plato’s forms; unreachable entities that exist beyond the descriptive world. Yet, as will be looked at, many, many moral theories are naturalist; they equate moral facts to some descriptive fact about the world, and so moral realism can be very much rooted in everyday experience. What moral realism fundamentally means more than anything else is that when we argue about whether something is right or wrong, we are having a substantive argument that can be reinforced by reference to real, factual statements. There are foundations to what we morally believe, and we are not simply vacuously arguing about something that is ultimately baseless. In fact many philosophers believe that moral realism is essential if we are to ever have real arguments about what is right or wrong, and that ideas such as moral progress only make sense if we are gaining a better understanding over time about what might be true and false moral statements.

However, there are also different forms of moral realism. Some philosophers believe that although there are objective moral facts, these can be never be written down as a set of rules; they are too complex, fuzzy, or unable to be grasped fully in language (this is known as moral anti-theory), while others believe that with a greater understanding of how morals relate to the natural world, we can come to increasingly better ways of listing moral rules, virtues or calculations. Therefore, there is a wide range of different opinions on moral realism, and by reading about them in the following pages you should try to judge for yourself which you think are the strongest.

Ethical Naturalism

So far, all the ethical theories we have studied assume that ethical language makes claims about mind-independent reality, hence they hold that **cognitivism** is true. Furthermore, utilitarianism makes the claim that ethical terms such as ‘good’ and ‘bad’ are associated with the natural properties of pleasure and pain. This meta-ethical position is known as **ethical naturalism**.

Ethical naturalism is a cognitive theory. It is the belief that decisions about what is right and wrong can be arrived at through discovery of the natural world and human nature. An action can be right or wrong if it fulfils the intended purpose of human nature (natural moral) or if it produces happiness (utilitarianism). Human nature and happiness are both things that exist in the world and can be assessed and empirically tested.

Ethical naturalism treats ethical statements the same as non-ethical statements. Just as we would expect the statement ‘that car is blue’ to be empirically tested (by looking at the car, I can see and be certain that it is blue), so too such ethical statements, such as ‘happiness is good’, can be empirically tested by analysing whether an action has produced happiness. On this basis, ethical statements can be proved true or false – they are verifiable or falsifiable.

Ethical naturalism developed as a response to the challenge of science to religious belief. It tried to show that ethics was as provable and definable as scientific facts.

We’ve seen this approach so far in utilitarianism and virtue ethics (whether Kant has naturalist or non-naturalist leanings is still a matter of debate), where what is good is equated with pleasure, or in excelling in the functions that define a human being. In fact many would argue that naturalism is simply a common-sense approach, for there are many things in the world that we consider good, so why would there not be a link between what is good and some feature of the natural world? Well, as explored, there is one prominent objection to naturalism that is still debated today.



#### Issues for Naturalism: The Naturalistic Fallacy

Critics of naturalistic methods of deriving morality are keen to point out these theories all make a simple, yet utterly damning, mistake: they confuse what *is* the case with how it *ought* to be. This is known as the is–ought gap, or the naturalistic fallacy. It was first pointed out by Hume (whose views we will discuss in more depth later):

In every system of morality, which I have hitherto met with, I have always remarked, that the author proceeds for some time in the ordinary ways of reasoning… when all of a sudden I am surprised to find, that instead of the usual copulations of propositions, is, and is not, I meet with no proposition that is not connected with an ought, or an ought not. This change is imperceptible; but is however, of the last consequence. For as this ought, or ought not, expresses some new relation or affirmation, ‘tis necessary that it should be observed and explained; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it.

- David Hume, *A Treatise on Human Nature*, Bk. 3, Part 1

In this oft-quoted passage, Hume is for the first time drawing a distinction which was to have a profound influence in twentieth-century meta-ethics. He makes the point that when philosophers talk about ethical matters they are prone to slipping from what *is the case* to what *ought to be the case* without ever explaining how they got from the former to the latter.

For example, I may claim that it *is the case* that everybody loves *Game of Thrones,* or at least everybody who has watched it. I may go on to explain, in depth, what it is about *Game of Thrones* everybody so much admires, how the universal love of *Game of Thrones* demonstrates its superiority over any other TV drama series, and how loving *Game of Thrones* has enriched my own and many of my acquaintances’ lives. I may then, quite naturally, conclude that everybody *ought* to love *Game of Thrones*. For Hume, it is at this point that the record screeches to an awkward halt. ‘At no point’, he might argue, ‘have you justified moving from the claim that it *is* the case that everybody loves *Game of Thrones*, to the claim that everybody *ought* to love *Game of Thrones*!’

Furthermore, Hume thinks it is unclear precisely how somebody could justify such a move. Even if it were the case that everybody on the entire planet loved *Game of Thrones* (which is, of course, not true – many people are indifferent to, or even hate, television drama, for instance) it is hard to see how this would mean they *ought* to. This is because to say that ‘it is the case that everybody loves *Game of Thrones*’ is to make a *descriptive* claim. It describes a factual state of affairs. ‘Everybody seeks pleasure and avoids pain’, for instance, is a descriptive claim like that made by Jeremy Bentham in the opening line of his *Introduction to the Principles of Morals and Legislation* (‘Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure’*)*.* Bentham would argue that this is a factual state of affairs and, I imagine, many of us might agree.

However, for Hume, Bentham makes a damning error in only the second sentence of the same work:

It is for them [pain and pleasure] alone to point out what we ought to do, as well as to determine what we shall do.

- Bentham, *Principles of Morals and Legislation* (my brackets)

This sentence is a *normative* claim. It is saying what we *ought* to do. That is a very different thing to saying how things are. It might have been the case that the vast majority of Afro-Americans were held in slavery in mid eighteenth-century Mississippi, but few today would argue it *ought* to have been the case. In fact, we might argue that it *ought not* to have been the case. Yet how can we decide who is right if the entire discussion is founded on a fallacy? The problem of whether it is possible to talk, assuredly, about how things *are* in the same way we talk about how they *ought* to be will occupy us for much of this topic.

#### Can Naturalism Move Past the Is–Ought Gap?

It can be too easily thought that naturalistic ethical theories are baseless at heart. If there is no way of deriving normative ethics from descriptive claims, then how can ethical statements be justified at all?

In truth, many philosophers question the pure distinction between is and ought statements, often arguing that the seeming differences between them are exacerbated by grammatical and semantic differences in the way we form moral statements and ordinary descriptive statements. So, if we look briefly at two of the main normative ethical theories we have studied so far, we can see how there may be ways to move past the is–ought gap.

#### Utilitarianism

If you remember, Mill himself stated that he could logically prove utilitarianism as the correct theory, but thought at least he could give a number of reasonable premises that, if accepted, could lead to a general belief in the principle of utility. Nevertheless, it was also noted that it does fall guilty to the fallacy of equivocation, and, as such, the naturalistic fallacy is a direct challenge to its central line of argument. So how can the utilitarian respond?

The main thing to note first is that utilitarianism is a teleological ethical theory. This means that all ethical action is directed towards an end; the greatest amount of happiness for the greatest number. What this also means, however, is that we can view utilitarianism as entailing a distinct set of hypothetical imperatives, such that any ethical statement can be formulated as:

‘If you want the greatest happiness of the greatest number, you ought to do…’

Therefore, there is potentially a descriptive element to utilitarianism. So long as we accept Mill’s premise that we naturally desire happiness as human beings, there perhaps is a valid move in formulating ethical statements as hypothetical imperatives towards this natural end. Much has been written about whether this is really possible, and it is important to note that one can refute the basis of utilitarianism by denying that the central aim of human beings is to seek happiness, but altogether, although the is–ought gap is problematic for utilitarianism, many still contend that one can reach normative, utilitarian statements from a set of descriptive conditions about the world.

#### Virtue Ethics

The same is potentially true for virtue ethics. For we noted that, in particular, virtues are ‘thick’ ethical concepts, which means they have both a potentially descriptive element, and a normative element also. When I call someone dishonest, I am expressing moral disapproval, but also identifying behaviour they have engaged in which I would class as deliberating choosing not to tell the truth. Furthermore, virtue ethics is also resolutely teleological, but, instead of happiness, the end is Eudaimonia. This means we can formulate similar ethical statements to utilitarianism. For example:

‘If you want to achieve Eudaimonia, you ought to do…’

What these two examples show in particular is that although the is–ought gap claims there is a fundamental distinction between the descriptive and normative statements, such that there is no logical connection between the two, it is possible to argue that this claim is missing the point. Whether or not we can divide between two different classes of things does not necessarily mean that the two things are not actually bound up in an important way. What this means is that it is important not to simply take the is–ought gap as a given; philosophers still debate today to what extent it undermines the claims that ethical statements make, and whether we can ever equate them with descriptive statements in any meaningful way.

Moore’s Intuitionism

G E Moore (1873–1958) is often associated with his Cambridge colleagues Bertrand Russell and Ludwig Wittgenstein. The work of these philosophers was considered radical at the time due to its focus on close linguistic analysis and a lack of interest in ‘speculation’. Moore, and his successors, all fundamentally felt that many of the puzzles philosophers had studied for millennia were actually not puzzles at all, but rather the result of confused language. They were extremely influential on what has come to be called ‘analytic philosophy’, which is the dominant way of doing philosophy to this day in English-speaking countries.

**Anthology Text:**

Moore, G E (1903),   
*Principia Ethica*

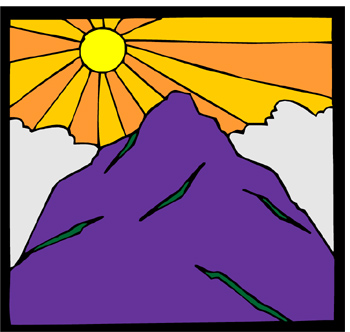


One of Moore’s most famous contributions is the ‘open question argument’. Moore considered an open question to be one that can be answered positively or negatively without ‘self-contradiction or betrayal of conceptual confusion’.[[3]](#footnote-3)

Moore’s example used the naturalistic element ‘pleasure’ (from utilitarianism), although other terms could be used. Utilitarians might say that ‘helping people who are ill creates pleasure’; therefore, ‘helping people who are ill’ is good.

Where the open question comes in is in asking in response to the claim of utilitarians, ‘Are pleasureful things good?’ The answer to this question could, without self-contradiction and logically, be that pleasureful things are not always good. From this Moore argued that ‘good’ is, therefore, not analytically (by definition) the same as ‘pleasure’. Therefore, ‘good’ is not the same as ‘pleasure’ and ‘good’ cannot be defined as ‘pleasure’.

A definition of ‘good’, therefore, has to be things that are analytically identical, e.g. things that are analytically identical to the term ‘bachelor’ include ‘man’ and ‘not married’. To ask ‘can a bachelor not be a man?’ is not an open-question argument because it would not make sense to answer ‘yes’. Therefore, we can be certain that bachelor means ‘man’ and ‘not married’. Moore argues that no such example using naturalistic terms could ever apply to the term ‘good’ because it will always result in an open question.



As a cognitivist, Moore did not, however, argue that this meant good did not exist. What is ‘good’ could be known and identified through our intuition – we know good when we see it – things are self-evidently good.

Moore uses the example of yellow to help explain this. Yellow is something that cannot be defined but we know it when we see it. We cannot say what yellow actually is – we can only say things that have yellow as a property, e.g. the Sun. We are, however, unlikely to dispute whether something is yellow or not.

We know what yellow is and can recognise it where it is seen, but we cannot actually define yellow. In the same way we know what good is. But that we cannot actually define it.[[4]](#footnote-4)

What is good is, therefore, known intuitively, and without recourse to any definable thing in the natural world. In fact, what is good is thoroughly indefinable as a non-natural concept, and any philosopher, such as Mill, who attempts to define it is engaged in a misguided project. Moore’s position is, therefore, called ‘intuitionism’, as opposed to naturalism. It is still a realist position, although one with a very different perspective on the status and nature of morality.

#### Moore’s Consequentialism

It may be tempting at this point to declare Moore and Mill as fundamentally opposed in their conception of ethics, but in some places they do agree. The main one is that Moore, like Mill, is a consequentialist. However, he does not argue that it is possible to in some way calculate what is the exact greatest good of one action over another, for the good is indefinable, and if it is indefinable it is also in some way not able to be quantified.

So how does Moore conceive of a good moral action? Well, he argues that in any ethical situation we should seek to promote ‘intrinsic goods’ through our actions. These are qualities, or ideals, that embody the intuitive goods that we can understand to simply be good in and of themselves. This may be the moral good, but it also includes concepts such as beauty and friendship. In this way, Moore is often called an ‘ideal utilitarian’, for he promotes the maximisation of these ideals whenever possible. However, his view is far from systematic, and many have accused it of being underdeveloped, with his brand of intuitionism often being the focus of criticism from emotivist philosophers, who will be examined later.

#### Issues for Intuitionism

Moore claims that we are able to intuitively recognise the property of ‘goodness’, but he does not explain how we do this or what exactly the nature of goodness is, only that it is non-natural and so unlike anything else.

Warnock points out that this raises a number of serious difficulties. Firstly, how do the non-natural moral properties interact with the natural properties? A utilitarian might say that poking people in the eye with needles is wrong because it causes them pain, but this route is not open to the intuitionist due to the open-question argument. Surely, however, that an action might cause suffering has some influence on our moral thinking.

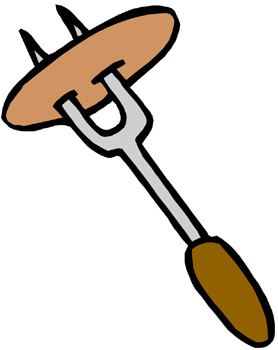
Secondly, we might ask: how does intuitionism explain moral disagreement? How does it aim to resolve ethical dilemmas? As a cognitivist, Moore believes there are moral facts and moral falsehoods, yet how do we know when we have found them? If I think abortion is always wrong and you think it is always the woman’s choice, how do we know who is grasping the moral fact and who the moral falsehood?

It gets worse; even if there are moral facts, the intuitionist gives no reason why anyone should care. If I judge that ‘torturing innocent children is wrong’ I may well be stating just such a fact, but simply stating it as a fact does not give anybody any reason *not* to torture children; all I’ve done, in effect, is state the obvious, I have not made it relevant to how persons should act.

We will see that this final objection, known as the problem of moral motivation, does not only trouble intuitionism.

1. Hume’s Fork, and Verification of Ethical Statements

David Hume, in his 1740 work *A Treatise of Human Nature*, divides statements into two different classes: relations of ideas and matters of fact. The former are statements that can be known a priori, without reference to experience, and the latter are a posteriori, statements that can be known only by reference to experience of the natural world. This divide came to be known as Hume’s fork, and as an empiricist Hume believed that relations of ideas could not convey real knowledge about the world; they only concern abstract reasoning between different ideas that we already possessed from experience.

One philosopher in the early twentieth century who was particularly influenced by this distinction was A J Ayer, and he in his 1936 book *Language, Truth and Logic*, outlined a position which, along with other philosophers’ views, came to be known as ‘logical positivism’. Ayer proposed that Hume’s fork provided a guide to what could be considered a meaningful statement. Unless a particular statement can be verified, either through it being true by definition or through sense experience, then it cannot be meaningful, as there is no real way of determining whether it is true or false. This test was called ‘the verification principle’; the logical positivists as a whole looked to apply it to different philosophical fields and determine whether their conclusions could truly be regarded as meaningful.

So how does the verification principle apply to moral statements? Well, they certainly aren’t true by definition, or we would all generally agree on what was morally good. Yet they also aren’t really empirically testable either. As we have noted, normative statements are different from descriptive statements, such that on making a statement such as ‘murder is wrong’, I cannot through my experience alone come to the conclusion that this is true or false. There is nothing in the act of murder itself that suggests that by itself as an act it is right or wrong.

Therefore, Ayer concludes that ultimately moral statements are meaningless. Though they appear to be able to be described as true or false, this misconception arises out of an error in our thinking. In reality, moral statements can never really be true or false, as they do not represent real propositions about the outside world. The results of this line of criticism, as well as responses to it, will be analysed in the section on emotivism; but one can still ask for the moment, is the verification principle a real test of meaningfulness?

2. The Issue of Moral Motivation

Another issue that Hume raised was that moral statements motivate us to act, while ordinary descriptive beliefs do not, meaning they cannot be the same thing, nor can moral statements be founded on ordinary beliefs or reason. For example, the statement ‘there is chocolate ice cream in the fridge’ does not motivate me to eat chocolate ice cream, for I may well actually intensely dislike chocolate ice cream. So the argument that ‘one ought to eat chocolate ice cream’ is not really based on any descriptive facts about chocolate ice cream, as the simple belief about its existence at a certain place or certain time does not contain a motivating argument for me to eat it.

But let’s unpick this idea of Hume’s for a moment. Firstly, we can say there is an argument about the categorical foundations of morality – simply that from its motivational content, it cannot be said to be the same as ordinary descriptions about the world. Secondly, a more nuanced point is that moral statements are different in that psychologically they are motivating, whereas ordinary descriptive statements are not psychologically motivating. No amount of description alone about chocolate ice cream makes me want to eat it. So Hume concludes that reason and belief are not the foundations of morality, rather it is our emotions and passions, the way we are disposed towards certain things in the world. So, when we make a moral statement, we are expressing our emotional state or approval towards a particular thing. This undermines the claims of moral realists, who want to say moral statements reflect some true state about the world. If they turn out to just be statements of emotions, how can we believe they have any objective content?

3. Moral Disagreement

It was noted in the introduction that disagreement is not necessarily a measure about whether something is real or true, but equally it can be the case that radical disagreement about a particular thing can indicate that it is ultimately illusory. This is especially true when the kind of statements under discussion in ethics are already abstract, and unable to be related to direct description about the world. This is partly the approach J L Mackie takes in criticising moral realism, for it is not just disagreement that reveals the lack of foundations of morality, but disagreement about a set of statements that are fundamentally different to descriptive statements.

Building off the criticisms we have analysed so far, Mackie argues that it is naïve to believe that our ideas about morality actually refer to anything real in the natural world, and the only way that progression in ethical discussion can happen is if we begin to understand the errors we have made in believing that there are objective moral facts. He begins by noting that across the world, different societies and cultures have had vastly different opinions about what is moral and good, such that there have been irreconcilable disagreements about these opinions throughout history.

Now, he notes that when we have such disagreements, there are two broad ways of understanding them. The first is that the disagreement is about something real, with our perceptions about this real thing simply being distorted, or uninformed about the objective reality under question. This means there is some truth underneath our disagreement, but we just haven’t discovered it. However, the other way of understanding disagreement is that there is nothing real underneath, and as such there is no objective truth that can be discovered or drawn upon. For example, say two people take hallucinogenic drugs and both see a leprechaun dancing in the middle of the street. One cries out that the leprechaun’s outfit is green and the other claims it is blue, and they both begin fighting in the middle of the street about the colour of the leprechaun’s outfit. In this case there is no real answer to the colour of the leprechaun’s outfit because the leprechaun isn’t even real in the first place.

Mackie wants to argue that moral statements are similar to the example of the leprechaun. Although we think we are having a genuine resolvable discussion when we talk about morality, we are really trading arguments about issues that have no objective truth behind them. In fact, anthropological studies might even be used to reinforce Mackie’s claim here, as vastly conflicting ideas about what is moral and good have been observed in different societies at different levels of development and technological progress. To this end it is possible to say that there is no unifying real basis to morality and so no real objective truth to moral statements as a whole.

Yet Mackie cannot make this argument alone. For the realist’s response is to simply claim that Mackie has made a mistake, and actually moral statements do refer to something objective and real underneath. Our capacities to debate moral issues, understand how the morals of different societies might vary, and arrive at moral compromises indicate there is as much potential unity as disagreement. Therefore, Mackie needs to provide a second part to the argument: why the disagreement arises in the first place.

This, he simply argues, is that in the case of morality, it is much more likely that disagreement arises as a result of there being no objective moral facts. When we compare ethics to science, although many people in the latter discipline have disagreed about the nature of the world and been wrong about their assertions, there still is a trend, using the scientific method, to discern real and important facts about the world which individuals can agree on. Whereas if we take monogamy (Mackie’s example), there is nothing particularly scientific that could resolve whether it is right or not. Some societies believe monogamy is important, others don’t, and the question of whether it is good or bad has not got any closer to being solved throughout the history of humankind. At this juncture, we can still uphold that at some point we might discover whether monogamy is right, but this is much more shaky as time goes on, our philosophical discussions become circular and the question remains unanswered. In this way, it is better to simply say that whether monogamy is regarded as good or bad arises out of traditions in different cultures, and there are no true moral facts about whether monogamy is good or bad.

Therefore, Mackie provides a secondary argument to indicate why moral disagreement potentially reveals the lack of objective truth behind ethical statements. As we will see, this line of criticism leads him to develop his own theory of meta-ethics, commonly called ‘error theory’.

#### Discussion:

*To what extent do you believe different societies and cultures fundamentally disagree about moral laws   
and principles? Are there any broad areas of agreement about what is right for all human beings?*

Moral Anti-realism

So far in this companion, we have only studied realist theories, and their associated criticisms. However, for many philosophers these criticisms lead to an obvious conclusion: there simply aren’t any objective moral facts or properties! In this section, two different forms of anti-realism will be examined. First, Mackie’s error theory, which is a cognitivist form of moral anti-realism, and second, emotivism/prescriptivism, which are more common non-cognitivist theories of moral anti-realism. In each case, anti-realist theories make some pretty radical claims about the nature of morality, and their arguments potentially have great ramifications for normative ethics.

Error Theory

Error theory is often associated with the Australian philosopher J L Mackie (1917–1981). Unlike other cognitivist accounts of ethical language, error theory does not seek to defend the idea that moral judgements make *true* claims about mind-independent reality. Rather, Mackie takes the peculiar position of claiming that ethical language *does* make claims about reality but that they are always *false* (thus, an error). Therefore, Mackie is an anti-realist, and affirms there are no mind-independent moral properties or facts.

**Anthology Text:**

Mackie, John L (1977), *Ethics*: *Inventing Right and Wrong*, Penguin, Chapter 1, Sections 8 and 9



Central to this line of thought is the *argument from queerness* (so-called because it states moral properties would be queer things indeed were they to actually exist). The idea is that most of our judgements – for instance, ‘it is 23 °C in Bristol this afternoon’ – refer to things which actually exist, as Mackie puts it, in ‘the fabric of the world’. Moral judgements have a similar form, e.g. ‘It is wrong to torture helpless animals’, but they do not seem to refer to anything in the world. Where in the action of torturing helpless animals can ‘wrong’ be found? Is it a part of the pliers used to tear off the flesh? Is it in the animal’s desperate yelps? Is it in the arm muscles with which the torturer swiftly jabs pins into the poor creature’s eyes? In contrast, it is easy to establish where ‘in the fabric of the world’ the temperature or the city of Bristol exists. These things are easily knowable; I need only look at a street sign and a thermometer to be sure my judgement ‘It is 23 °C in Bristol this afternoon’ is correct.

Mackie is in fact making two distinct but related claims about moral properties:

(i) An *epistemological* claim (about what we can know): we cannot know about moral properties (or the connection that holds between natural properties such as pleasure and moral properties such as goodness).

(ii) A *metaphysical* claim (about what exists): moral properties are not a part of mind-independent reality.

The justification for (i) is that, unlike natural properties (weight, temperature, etc.), there is no way for us to verify that moral properties exist.

The justification for (ii) is that moral properties, were they to exist, would be unlike anything else in the known universe. This is because moral properties *motivate* action. My telling you that if you torture small animals you are doing wrong may motivate you to stop, particularly when I tell you it is one of the earliest visible manifestations of psychopathy. There is no other type of thing which has this peculiar power. In fact, it is impossible to see how any other part of reality *could* have this power.

#### Morality and Psychology

There is still one thing to be explained in the error theorist’s discussion of morality. If there are no objective morals, why does it seem as if there are? Why are our moral beliefs so important and forceful? Mackie argues that there is a psychological explanation for this, namely in the way we are conditioned to think of things as good or bad from our upbringing and our place in the communities we live in.

In every society there are often arrangements, traditions and rules by which people are bound, often to secure the future stability and survival of people. However, these customs or rules over time are falsely objectified by people within the culture, even though they have no objective basis. This means that people end up making an error by supposing their moral statements have truth to them, when in fact they are just reflections of custom, tradition and culture.

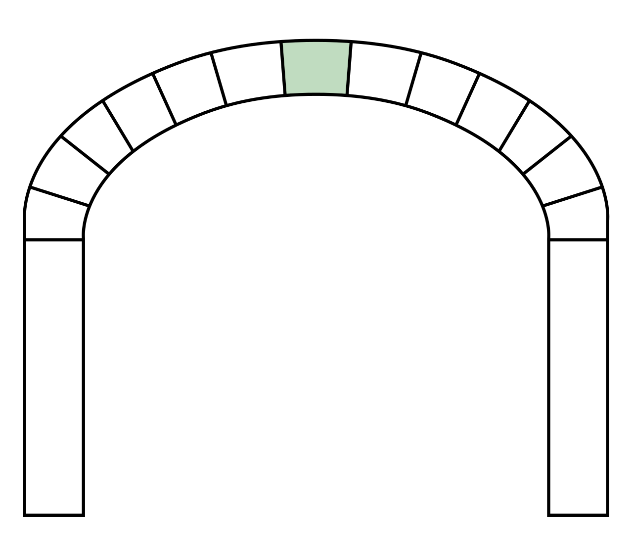
#### Issues for Error Theory

The central claim of Mackie’s error theory is that moral properties seem unable to fit into our epistemological and metaphysical framework. However, it might be argued by some of the thinkers we have studied that there is nothing particularly ‘queer’ about moral properties properly considered. For Kant, good actions are those which are done in accord with reason, suggesting that the moral property of ‘goodness’ is an intrinsic part of reason itself. We know something is good because it corresponds to some feature of reason (i.e. it is universalisable or treats rationality as an end in itself). Likewise, a utilitarian may wish to suggest that the moral property of goodness simply exists in the experience of pleasure, and the moral property of wrongness exists in the experience of pain.

#### Are Our Moral Beliefs Really That Different?

There are two important ways to approach this question. One which was examined previously is simply to say moral statements aren’t really that queer. Mackie is just mistaking semantic difference in the way we phrase facts about the world with some genuine distinction between moral and descriptive statements. Many point out in particular that so long as morality is perceived to have some teleological end then moral statements really aren’t that queer, they just refer to human action against teleological aims. In fact, this potentially leads on to the second issue; whether human morality is really that different culture to culture.

This is a much disputed and hard to quantify issue. What looks like difference to one person, may seem remarkably similar to another. Some critics have contended that there are actually broad areas of agreement in moral issues between different societies and cultures. For example, the golden rule, ‘do unto others as you would have them do unto you’ is an intuitively understandable principle observed in many different societies, and in itself only applies to a broad understanding of personhood, not cultural or social traditions. Mackie, in using the example of monogamy, is perhaps obscuring genuine agreement by picking a custom of which there is less a moral basis, and more of a cultural one. It might also be possible to state that even if there are great areas of moral disagreement, most people aim for similar moral ends, such as happiness or flourishing. If there are agreements in the teleology of ethics, then arguably it is possible to have effective realist discussions about what is right and what is wrong. Therefore, Mackie’s case is less than clear-cut. The queerness of morality may be overemphasised, along with the moral differences between cultures.



***Moral Motivation***

The problem of moral motivation is an issue for cognitivist theories generally and is error theory’s main   
line of attack. To successfully defend themselves, cognitivist theorists can do one of two things:

i) They can simply deny morality is motivating. Morality does not motivate all people; it is certain   
that there are individuals past and present who are fundamentally amoral – good, bad, right and wrong mean nothing to them. Goodness may be a feature of the ‘fabric of the universe’ that we can discover, but the amoralist has no interest in acting upon such knowledge.

ii) They can deny that ethical language relates to anything in the natural world. Instead, they can say that ethics talks about *reasons*. If we consider something to be a reason, then necessarily it is a reason for a person to believe (or not to believe) something. For instance, finding a bloodied knife beside a mutilated corpse is a reason to believe a murder has taken place. This is a fact – it just *is* a reason to believe that. It doesn’t mean a murder has *actually* taken place; it could be an elaborate hoax, a suicide, a bizarre accident.

In the case of ethics, we would have to reflect on the reasons we have for certain moral beliefs. I believe rape is wrong. What are my reasons for this? Because it causes so much hurt and damage. What reason do you have for not desiring hurt and damage? It does not bring happiness. Why do you desire happiness? We could go on like this. Yet, the argument goes; this does not mean the endeavour is futile. Rather, it means we must accept that our reasons justify one another as each piece of an arched bridge supports the other.

Remove just one and the entire structure collapses. As rational beings, reasons necessarily motivate us – we cannot rationally ignore reasons. We decide what is right – objectively – by removing those reasons which make no sense, and by finding those which do. When we do ignore reasons, we are not being rational.



**Stretch and Challenge**

Ethical Language 2: Non-cognitivism

Ethical non-cognitivism refers to the belief that moral statements are not statements that can be considered to be true or false in reference to empirical evidence or the natural world. They are statements that do not attempt to express genuine propositions or assertions that describe some moral reality. Therefore, non-cognitivists are moral anti-realists, and do not believe there are any mind-independent moral properties or facts.

An example of a non-cognitive statement is: ‘I like honesty’. This statement refers to my subjective, emotional state. It does not claim to describe anything that exists in reality or is objectively true. It does not make sense to say ‘I like honesty is false’. It is not fitting to use the language of ‘true’ and ‘false’ in this context. This does not necessarily mean the statement ‘I like honesty’ is completely meaningless – it does achieve certain other purposes, such as telling other people what I like or suggesting to others that they too should be honest.

An example of a similar cognitive statement is: ‘He is an honest man’. This refers to a certain type of behaviour which could be proved true or false in reference to the empirical world, i.e. whether the man tells lies or not.

Why instinctively be an emotivist, however? This will be explored throughout this section, but one major reason is that although one might want to explicitly deny moral realism, there are equally difficult reasons to suppose that theories such as Mackie’s error theory can effectively challenge the claims of moral realism. It may be that the queerness of moral statements reveals that they are not actually real statements in the first place! The implications of this are great, although philosophers still disagree about whether the conclusions of emotivism warrant a complete re-evaluation of normative ethics, whereby we no longer talk about whether things are right or wrong in the manner we are used to.

Emotivism

Emotivist theories of ethical language focus on the emotive meaning of words such as ‘wrong’ and ‘good’. They want to argue that ethical language is essentially *emotive* and *expressionistic* in character – it expresses our feelings of approval or disapproval about an action or an individual.

There is an intuitive appeal to the idea. Ethical language often has an emotional pitch, ‘You no-good thief!’, ‘Murderer! Murderer!’, ‘The sheer wickedness of the Holocaust is palpable when one walks through the deserted corridors and windowless cells of the former concentration camp.’

The position can be traced back to David Hume, who first drew a distinction between descriptive and prescriptive statements:

1. Descriptive statements = ‘relations of ideas’ which are analytically true (e.g. all bachelors are men) and ‘matters of fact’ which can be empirically verified.

2. Prescriptive statements = what ‘ought’ to happen or be done.

**Anthology Text:**

Hume, D (1739–40), *Treatise of Human Nature*, Book III, Part 1



Here are some examples:

• Descriptive statement: Killing people causes harm.

• Prescriptive statement: It is wrong to kill people.

• Descriptive statement: I think you look nice in yellow clothes.

• Prescriptive statement: You should wear yellow clothes.

Hume argued that it is wrong to derive an ‘ought’ (a prescriptive statement) from an ‘is’ (a descriptive sentence). For each descriptive sentence above, it is wrong to derive the sentence directly below it. This critiqued ethical naturalism because it argued that from statements derived from observations of the natural world, it was impossible to make any ethical judgements. This is often known as the ‘is–ought controversy’ – that you cannot derive an ‘ought’ from an ‘is’. Hume, therefore, believed there could be no moral facts, but instead sentiment and feeling were the source of right and wrong.

Ayer’s Emotivism

You might have previously studied Ayer in Epistemology or Metaphysics of God. Ayer did not just turn his verification principle on to religious language, but on to ethical language too.

You will remember that Ayer claimed language was only meaningful insofar as it was either:

a) *Analytic* (that is, tautological, e.g. ‘All bachelors are unmarried men’)

**OR**

b) *Synthetic* (that is, subject to empirical observation or experimentation, e.g. ‘There are fourteen blue cars in the car park today’)

Ethical language is clearly not tautological (the concepts of murder and wrongness are not synonymous, for instance) nor is it possible to determine empirically whether a statement such as ‘Laughing at failed musicians is wrong’ is true or false. This is because goodness and wrongness are not things which are part of reality, so statements about them cannot be true or false.

**Anthology Text:**

Ayer, A J (1946), *Language, Truth and Logic*

Ayer, A J (1973/1991), *The Central Questions of Philosophy*



Instead, Ayer believes ethical language expresses our feelings about a certain course of action, behaviour or characteristic.

This has led to the theory being unflatteringly referred to as the ‘Hurrah! Boo!’ theory of morality, as it seems to reduce ethical statements such as ‘The mass murder of over two million people during the Cambodian genocide was an act of unspeakable evil’ to ‘Boo! The Cambodian genocide! Boo!   
Mass Murder!’

**Anthology Text:**

Warnock, G J (1967) *Contemporary Moral Philosophy, New Studies in Ethics*



#### C L Stevenson and Emotivism

C L Stevenson presents a slightly different form of emotivism from Ayer. He agrees with Ayer that moral statements are not real truth-evaluable propositions, but he disagrees that they are mere statements of feeling or emotion. In fact, he emphasises an important point of Ayers; that moral statements are intended to win over or influence people to a particular point of view. Leading on from this, the way in which we engage in moral discussion is not simply a trading of feelings, but a representation of our particular approval for a particular action or view. Therefore, when we make moral statements, we are referring to a certain sort of moral preference, which naturally has emotions, feelings and passions bound up within it. This view, Stevenson believes, more accurately reflects the way people not only held particular moral views, but also the way they used those moral views to persuade people to their way of thinking. Therefore, it is important when discussing emotivism to note the subtly different ways philosophers approach the way moral statements can potentially be reduced down to statements of emotion or approval.

#### Issues for Emotivism

In *Contemporary Moral Philosophy*, Warnock argues that one major issue with emotivist theories is that they fail to properly distinguish ethical language from other forms of language, which creates an emotional impact (for example, the impassioned speeches that are the hallmark of melodrama).

More significantly, it seems that ethical language is not even necessarily emotive in content. Sometimes, as we are doing now, ethics is discussed dryly, even matter-of-factly. Historically, ethically momentous decisions have been made not in floods of tears or with wringing hands, but through discreet memos in a bureaucrat’s hand.

You may also remember from previous studies a fatal flaw in Ayer’s reasoning: the verification principle itself is, in fact, meaningless by its own standards.

‘All meaningful statements are either synthetic or analytical’ is itself neither an analytic nor a synthetic statement!

We can, therefore, ask to what extent one might be able to identify moral statements as meaningful beyond any test of verification. This is something that will be explored in depth in the general issues of anti-realist views, but it is important to note that a large part of Ayer’s views require moral statements only having particular meaning on a subjective, emotional level. If we can possibly identify a way in which moral statements relate to a particular moral description of the world, beyond the states of feeling or emotion Ayer argues they should conform to, then it can be argued the central claim of the emotivists is inaccurate, or at least incomplete.

#### The Circularity of Emotivism

One criticism, made by Alasdair MacIntyre, is particularly pertinent. He asks that if we happen to accept the emotivists’ claims about moral statements, ‘what kind of feelings or emotions do moral statements really express?’ When we make moral statements we usually are, as Stevenson contends, making some statement of approval or disapproval about something, not just expressing an emotion. But we can press the question further and ask ‘What kind of approval are we expressing when we use moral statements?’ For clearly, when I state ‘Ice cream is good’ I am expressing something different to when I say ‘Genocide is good’. With the former I am not trying to persuade anyone of the general goodness of ice cream, nor am I saying people ought to eat ice cream. But if I say the radical and abhorrent moral claim ‘Genocide is good’, something more is being expressed in my approval of genocide; what we might call a moral element.

Therefore, we can identify many different kinds of approval. So MacIntyre claims when we ask ‘what kind of approval?’ the emotivist either does not specify anything which can differentiate the various forms of approval we use, or says that they are talking about ‘moral approval’. However, this latter option is very circular, for the moral realist would equally claim that in any moral claim there is moral approval; simply that this moral approval corresponds to some moral fact. Moral approval in this sense is agreeing with an act based upon the fact I think objectively it is morally good, not just because I have an emotional approval to that act. The emotivist, however, wants to contend that all that is moral is actually just unjustifiable emotional approval, and is still left having to explain what they mean by moral approval.

In short, MacIntyre contends that the emotivist cannot explain specifically why moral statements are unjustifiable states of emotional approval without reference to these particular states as being some form of ‘moral approval’, and this leads to circularity, as shown in the diagram below.

They are unjustifiable or unverifiable states of emotional approval.

How are these statements of approval different from others (e.g. prudential)?

They are unjustifiable statements of ‘moral approval’.

What make these statements of ‘moral approval’ and not just ordinary approval?

The emotivist ultimately wants to contend that moral statements are special or different in some kind of way, such that we can reduce them to emotional states of approval, but at the same time struggles to find any way to define this specialness without referring to the ‘moral’ aspect of moral statements. This suggests that moral statements do have a particular form of meaning beyond their emotional content, and that how we use moral statements is not the same as how we show feelings of approval or disapproval towards different acts and ideas. MacIntyre’s point is a subtle one, but it shows how, although one might initially be able to equate moral statements with emotional statements, this equivalence is not straightforward and can easily be vacuously circular.

How else might we, therefore, understand moral statements non-cognitively? One solution, prescriptivism, will be analysed in the next section.

Prescriptivism

#### Richard Hare (1919–2002)

The philosopher Hare, in his books *The Language of Morals* (1952) and *Freedom and Reason* (1963), developed the ethical theory of prescriptivism.

Prescriptivism understands moral statements to be prescriptive – telling or prescribing what someone ought to do. It is similar to a doctor’s prescriptions, which tell patients what medicine or exercise they should take.

People, therefore, make moral statements as a way to convince others or influence them about what they should do. If I say ‘murder is wrong’, I am saying ‘you should not murder’, with the hope that my listeners will respond to what I have said by not committing murder. Hare writes:

**Anthology Text:**

Hare, R M (1952), *The Language of Morals*



I have said that the primary function of the word ‘good’ is to commend. We have, therefore, to inquire what commending is. When we commend or condemn anything, it is always in order, at least indirectly, to guide choices, our own or other people’s, now or in the future.[[5]](#footnote-5)

An important feature of prescriptivism is that it states that moral commands should be **universalisable** – they should apply to everyone. Therefore, if I say ‘murder is wrong’, this would mean ‘everyone should not murder’. The prescriptions should be consistent (mean murder is wrong in all situations) and not be hypocritical (it is not right for other people to murder, but it is right when I do it).

#### The Advantages of Prescriptivism

We’ve explored in part how emotivism seems to fall short of describing what really makes up a moral statement, and one advantage of prescriptivism in particular is that it seems to capture more how we use moral statements as a guide for action. The universal element of a moral statement, the ‘ought’ or the ‘should’, is noted beyond the emotional element, and so, arguably, Hare does a better job of explaining how we rationally discuss ethical dilemmas beyond simply expressing our approval or disapproval of something. The ‘value terms’ involved in ethics, such as ‘good’ or ‘bad’, have the important function of commending action, and this is what prevents them from being effectively reduced to emotions or feelings.

How can we more concretely interpret this? Well, to take utilitarianism as an example, if I were to say ‘We want to achieve the greatest amount of happiness in all our actions’, then really I am recommending to someone ‘You ought to act so that your actions result in the greatest amount of happiness’. There is no logical connection between these two things. Hare himself believed that Moore’s open question argument refuted naturalistic moral theories, but equally our ethical propositions cannot be understood as anything but calls to action.

On a more everyday case, I might say ‘It is wrong to watch violent horror films’ to a friend, but what I am really expressing is ‘You should not watch violent horror films’. There is no moral fact expressed by this call to action, and it may well be that my friend ignores my call to action and goes and watches the latest film in the *Saw* franchise the next weekend. But what prescriptivism does capture here is that I am not simply expressing emotional approval of a certain action. Instead my call to action here may have reasons beyond that of emotion, and as such my friend and I could potentially have a more rational discussion around the issue than simply going ‘hurrah, boo!’

What is important to also note is that Hare does not also argue that all imperatives are moral ones, simply that moral statements are a form of imperative, namely universal imperatives. I might shout to my friend to drive on the right hand side of the road if we’ve gone abroad from the UK, but this equally does not apply when we are home. It’s a hypothetical imperative, whereas moral imperatives imply universality. If I shout a moral command to my friend, it equally applies to anyone if they are in the same situation as my friend.

Therefore, we can see a number of key potential advantages to prescriptivism:

1. It accommodates a rational basis to ethical discussion.

2. It identifies a common universal element to morality, present in how we use moral statements.

3. It notes how moral statements command others to action, not simply express our approval or disapproval of something, or influence others towards our views.

#### Issues for Prescriptivism

The main issues for prescriptivism are again to be found in Warnock’s text *Contemporary Moral Philosophy.* Warnock begins by outlining an interpretation of Hare’s theory:

*The thesis that moral judgments are prescriptive implies that one who accepts the moral judgment that he ought to do X is logically committed to doing X*

In Warnock’s view, Hare is taking a similar position to Kant on the way we issue moral judgements: they should be universalisable.

*that is, if I commit myself to this judgement in your particular case, I thereby commit myself to the view that anybody – including, most importantly myself, – in the circumstances in which you now are ought to act in that way.*

Warnock then proceeds to make two closely related criticisms of Hare’s account of ethical language:

i) ‘moral judgements are not *essentially* prescriptive’

ii) ‘if [they are not essentially prescriptive], we need not claim for ‘universalisability’ the importance which Hare, as I think mistakenly, claims for it.’

i) As he did to counter the claims of the emotivist, Warnock argues that moral language has multiple uses: *‘They may be prescribing, certainly; but also they may be advising, exhorting, imploring; commanding, condemning, deploring; resolving, confessing, undertaking; and so on, and so on.’*

Hare restricts himself to a very limited context (those in which imperatives are issued by a speaker) and then, incorrectly, declares this is the sole realm of moral discourse. The central doctrine in Hare’s account is that ‘there obtains a quite special connection between words and deeds’. Namely, to accept or reject an ethical statement, for Hare, is just to act or not act on it.

However, Warnock believes that while it is true that imperative discourse has a single, clear relation to conduct (you either assent and follow the order, or dissent and disobey), moral discourse does not.

When it comes to moral discourse *‘someone’s conduct will be somehow related to ... what the speaker says ... But the actual relations, quite clearly, will be wildly diverse, and not to be summed up in any single   
formula whatever.’*

ii) In Hare’s view, to issue a prescriptive statement is to be willing to universalise it. So far, so Kantian. The problem is that Hare does not believe there is any objective standard of goodness – that is why he is considered a non-cognitivist. Kant does, however; he thinks the only good is the good will, and that is determined by reason. There is no only good for Hare because that would involve saying ‘It is *true* that x is good’ or ‘It is *false* that x is good’.

Without such a standard, Hare can only demand that a person is *consistent* in their judgements. So, if you believe that sexism is ethically right, so long as you are consistently misogynistic, and consistently prescribe that people treat one gender as superior to the other, then we can have no grounds on which we may rationally disagree. So, Warnock concludes, Hare’s attempt to rescue cogent moral argument is *‘as hopeless as trying to play a competitive game in which each competitor was making up his own rules as he went along’*. Since there is no independent moral standard from which to judge our prescriptions, any person might be able to effectively and consistently universalise both trivial and deeply immoral prescriptions, without any way for us to rationally criticise them.

Issues for Moral Anti-realism

We’ve explored a number of individual ways in which both emotivism and prescriptivism fail to capture the way we use moral statements, but there are also more general criticisms of anti-realist positions. Some of these arise out of the specific criticisms made of both emotivism and prescriptivism, but some also arise out of the way we engage in normative ethics, and the way we might be able to perceive the development of our morals throughout history. Therefore, while it is useful to know why emotivism and prescriptivism might fail in themselves, it is also useful to be able to justify why one might hold moral realist views above those of the anti-realist.

1. Moral Progress

It is not an uncommon view to hear those who have just started analysing ethics to declare, ‘well there’s so much disagreement primarily because morality is just subjective’. However, such a view is not only much more difficult to uphold coherently, it also is without justification, a rather trivial analysis of the way we view the role of ethics in our lives. Moreover, very few people in the world are willing to abandon all ethical principles, especially if they are on the receiving end of someone else’s vindictiveness or malice. So care needs to be paid to not only the way morality is employed by human beings, but the way it is potentially possible to view ways in which human beings may have morally developed or grown in their beliefs over the course of history. For if morality did develop throughout history, then ironically we would expect to see moral disagreement between different societies, people and cultures, who may well be at different levels of moral growth. Just because there may be moral growth does not mean that everyone in the world would advance or grow at the same time.

So why should we believe there is a moral growth, and why would moral growth pose a challenge to the anti-realist? The simple argument for the former is that we reject practices of our ancestors, and we do so not because those practices would not be beneficial towards us, but because we regard such practices as morally wrong. To take the example of slavery, there are still many people who illegally enslave others, and, even if these slavers offered individuals the chance to own slaves in the UK, it is probable most would turn such an opportunity down, even if it potentially would bring a financial benefit.



Now at this point the anti-realist might turn round and say, ‘well this is just an example of changing attitudes and emotions, not a reflection of moral growth’, but such a view is still difficult to completely uphold. If we think about the United States Declaration of Independence, you might know of the famous phrase within it that guarantees as rights to all people in the US ‘Life, Liberty and the Pursuit of Happiness’. Now, despite this clause being written in 1776, it almost certainly wasn’t applied to everyone in the US. Slavery was still legal, and systems of economic distribution were not set up such that everyone could be free and pursue happiness. So we might say that while the founding fathers had moral ideals, they did not yet realise what those moral ideals meant, and that, over time, as understanding of these concepts grew, it was realised that it was not possible to hold liberty as a right for all peoples while keeping slavery legal. Many philosophers argue here that the easiest way to explain such a change is that there are certain objective moral facts, and that throughout history moral growth has at times occurred such that the moral codes of societies are better and closer to these objective morals than they were previously.

This is far from an easy conclusion to draw, however. Even if one does not accept the view that there are no moral facts, one may still well hold up moral relativism, arguing that this moral growth we observe is only relative to a particular culture. The abolition of slavery in the US as such can only be understood as moral growth relative to those important moral ideals that were already part of the culture of the US, and make little sense if we compare them to other cultures which hold vastly different ideals. However, there are a number of responses to this issue also.

The first is that it supposes in a sense that cultures have never been influenced by each other in their ideals. Throughout history different societies have traded in ideas and have been perceived to have grown in different ways from the influence of other traditions and cultures. For example, Christianity was influenced not just by the words of Jesus, but also Jewish traditions, Greek philosophy and many other separate smaller cultures that adopted it as their official religion over the years. In this way, it might not be so easy to simply relate ideals to a particular culture, for no culture is really static, and many evolve rationally in response to ideas and thoughts introduced by other cultures around them.

Secondly, an important point to make is that simply because there is moral growth does not mean that societies or people are always bound to morally grow. Instead this course may be interrupted, reversed, or never even start at all. History may well be full of periods where the moral lessons and facts learnt are reversed, and people become literally morally worse. What this means in particular is that there is no particular period or isolated part of human history which one can point to and definitively argue as demonstrative of the truth of anti-realism.

Overall, the issue of moral growth does not give wholesale support to moral realism, but it equally presents a difficulty for anti-realism. Beyond simply saying there is disagreement about morality, the anti-realist has to show why instances in which it seems societies and people have morally grown have some alternative explanations that better fit an anti-realist narrative. It may well be such a task is impossible, but it is a task that should not be ignored.

#### Discussion:

*Do you believe there has been moral progress in the world? If so, is this a sign that there are potentially   
objective moral facts?*

2. Anti-realism and Moral Nihilism

What is moral nihilism? It can best be characterised by a rejection of moral laws and values, and so also a rejection of their validity or importance in one’s life as a whole. However, it should also be distinguished from amorality, which simply refers to the state of not caring about morality. Nihilism, instead, is a much more severe position, denying that human beings should engage in moral practices at all. If this is true, we should no longer speak of things as morally good and bad, and not say that any particular action was morally right or wrong.

This is an extreme position to say the least. Nihilism itself is often bandied around when people believe that individuals or societies have stopped caring about morality as a whole, and is generally used pejoratively to describe an undesirable situation in the moral landscape of humanity as a whole. From the Marquis de Sade to the German ex-band Autobahn in the film *The Big Lebowski*, nihilism is often used to refer to those who are at peace with performing acts or holding beliefs we would consider to be morally unacceptable. As such, the charge that anti-realism descends into moral nihilism has to be taken seriously. If there are no objective moral facts, should we abandon the conventional ways we even engage in moral discourse?

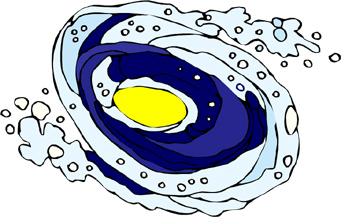
Some anti-realists claim that simply because there are no objective moral facts does not mean that we cannot continue our current ways of ethically discussing beliefs and facts, only that the way we approach ethical dilemmas requires a few tweaks. However, others go further, arguing for a middle way between moral nihilism and conventional ethics. This is what J L Mackie generally argues, contending that ethics has to be built up again from our understanding of error theory, such that we arrive at new ways and conventions of discussing morals as a whole. Anti-realism is, therefore, a reason to suppose that we have to initially adopt a somewhat nihilist perspective on morality, before rebuilding a vision of how ethics can survive and be useful in our everyday lives.

A similar view can be found in existentialist philosophers such as Jean-Paul Sartre. He proposed that ‘existence precedes essences’, meaning that all human beings are born ‘thrown’ into the world, with radical freedom to do whatever they wish. This means that, at heart, human beings are free from whatever moral obligations are thrust upon them in their upbringing, and they are free to develop their own meaning and morals in their life as they see best.

However, in many ways such positions seem difficult to argue for without acknowledging that it may be valid to simply do away with ethics altogether, and that if one embraces the radical proposition that there are no moral laws, one should not shy away from arguing it has no relevance to life whatsoever. In fact, many accuse existentialism of being nihilism in disguise, dressing up nihilism in an aesthetic or philosophical covering to make it more palatable.

Yet it should be noted that the charge of moral nihilism isn’t really a direct criticism of anti-realism. It simply states that the anti-realist can’t have it both ways. They can’t deny the existence of objective moral facts yet cling on to the conventional ways ethics is discussed and used in everyday life, and this may be enough of a reason to suppose that there is something at the core of ethics that is objective. For, if the very frame of human existence and interaction requires some binding ethical code, then it may be there is an objective element to morality that is built out of this requirement of human functioning or direction towards happiness.

3. The Distinctiveness and Use of Moral Language

This last thought leads perhaps to one of the most troublesome criticisms of anti-realism: the theories fail to describe what is actually going on when we use moral language. We’ve noted already within emotivism and prescriptivism that each potentially fails to account for particular ways in which moral statements hold meaning, and a key part of these failures perhaps can be traced back more generally to the tests of verification that underlie what counts as meaningful for the emotivist or prescriptivist.

For example, it has been noted that Ayer regarded the verification principle as the ultimate test for what is meaningful. Yet many statements we make which we would still count as meaningful don’t pass that test, from historical facts to everyday statements about art or film. If I stated the original *Star Wars* films are better than the prequels or newer sequels, although there is no way of empirically proving this, I could back my argument up with reasons we might still count as meaningful. I could say, ‘the cinematography is more attractive’ or ‘the storyline was more original’ and we would not simply disregard what I was saying based on the fact it was untestable, but instead view my assertions as perhaps a set of standards from which to judge other opinions against.

But what has this to do with morality? Well, one issue with emotivism in particular is that it presumes that moral statements gain their meaning from their relation to some real connection to some objective state of the descriptive world (which is supposed to be testable). However, this may not be the case. As Alasdair MacIntyre notes, what is really meaningful about moral statements is their relevancy and importance to a particular set of people, and their aims, beliefs and traditions. What makes a moral statement meaningful, therefore, is not whether it is descriptive of the world in general but whether it has certain imperatives that are important to a particular set of people. I might make the statement ‘people should always have a fire extinguisher in their house’ but if I live in a society that lives in tents, such a proposition would hardly make sense. Similarly, the phrase ‘one should not cheat on one’s spouse’ relies on having a pre-existing societal framework that encourages monogamous marriage.

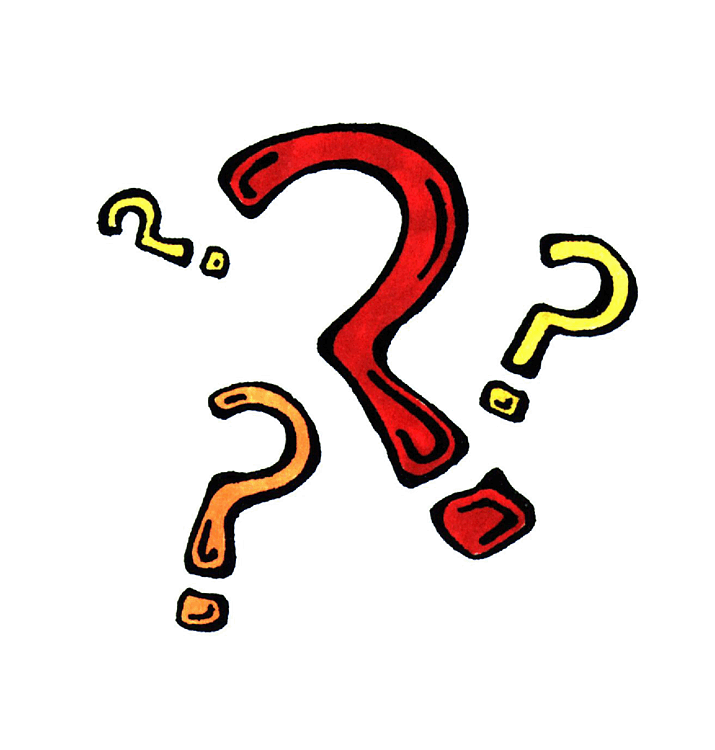
Now, MacIntyre believes moral statements can only be judged relative to the particular cultures and traditions they arose out of, but what his criticism demonstrates is that the anti-realist conflates meaningfulness with a relation to purely descriptive affairs. Therefore, we cannot discount morality simply because we cannot test it empirically in the same way that we do other propositions, just in the same way that we cannot empirically test my theories about *Star Wars*. There may be an objective fact that one set of *Star Wars* films is better than another, but the narrow criterion of meaning philosophers such as Ayer develop to determine whether this is true is not up to the task. An account of whether morality is meaningful and objective similarly has to relate to the way morality is used and the way we make moral judgements, and if a particular reductive account does not encapsulate these aspects then it arguably is criticising an incomplete view of ethics.

#### *Disagreement and Moral Ends*

This issue can be particularly seen if we look at cases of moral disagreement. If emotivism, for example, is correct, then it would seem as if there is no point in having a moral argument, for we would never really be disagreeing about anything concrete and simply just voicing our emotional approval in an attempt to sway the other person. Similarly, we would never reach a rational conclusion since there is no rational middle ground to be had, and there is no real fact of the matter to be discovered. Similarly, it was noted that despite prescriptivism’s attempt to solve such an issue, the lack of objective standards makes ethical discussion or argument equally problematic.

In fact, if a friend of yours exclaimed ‘eating babies is acceptable’ and you vehemently rejected such a proposition, you’re not really expressing an emotional state, but rather directly contradicting the rational chain of thought that leads to such a conclusion. In arguing with your friend you might appeal to numerous rational principles, such as the golden rule, or that eating babies does not lead to overall happiness. In short, you would offer reasons that would seek to shore up your argument as more rational and true than your friend’s, and so moral discussion is not really reducible down to emotion, nor is it literally talking about some description of the world. It is a set of rational arguments bound up in what it means to be moral, whether it be the laws we should have, or the moral ends we should aim for. In this way, it can be contended that both emotivism and prescriptivism mischaracterise in morality, and so claim that it is not objective based on a false set of criteria for what they suppose moral statements to be.

What this means is that one has to be very careful in outlining the foundations of ethics. If one is adopting an emotivist or prescriptivist position it is not simply enough to criticise certain aspects of moral statements, or reduce them in ways that don’t capture the full extent of how we engage in moral discussion. Doing so may be ignoring the fundamental aspects of ethical discussion that reveal how morality can be objective, and how there may be moral facts in the world.



**Meta-ethics: Quick Quiz**

1. What is cognitivism?

2. What is the naturalistic fallacy?

3. What is moral anti-realism?

4. What is the difference between analytic statements and synthetic statements?

5. What is prescriptivism?

Appendices

## Appendix 1: Online Resources for Teachers and Students

Go to **zzed.uk/9765**

🖰

***In Our Time* – Philosophy Archive**

*http://www.bbc.co.uk/programmes/p01f0vzr*

Practically a degree free of charge; most of the guests on these podcasts are leading authorities in their field and the discussion is almost always clear and to the point.

Recommended episodes: Mill, Logical Positivism, Just War, Utilitarianism, Relativism, Virtue, Duty, Happiness

**Philosophy Bites**

*https://philosophybites.com/*

Similar to *In Our Time* but focused solely on philosophy; also seems to cover recent research.

***Stanford Encyclopaedia of Philosophy***

[*http://plato.stanford.edu/*](http://plato.stanford.edu/)

Probably pitched at too high a level for all but the most capable students, but the SEP is still a fantastic reference tool.

***Internet Encyclopaedia of Philosophy***

[*http://www.iep.utm.edu/*](http://www.iep.utm.edu/)

Often a slightly easier read than the *Stanford Encyclopaedia* in my opinion. Covers some topics Stanford doesn’t, and vice versa.

**8-Bit Philosophy**

[*https://www.youtube.com/playlist?list=PLghL9V9QTN0jve4SE0fs33K1VEoXyL-Mn*](https://www.youtube.com/playlist?list=PLghL9V9QTN0jve4SE0fs33K1VEoXyL-Mn)

A mash-up of old Nintendo games with even older philosophy.

**A History of Ideas**

[*http://www.bbc.co.uk/programmes/articles/4FYLvwSKzJ5vwNv9Z9mwbSD/a-history-of-ideas*](http://www.bbc.co.uk/programmes/articles/4FYLvwSKzJ5vwNv9Z9mwbSD/a-history-of-ideas)

A series of short animations commissioned by the BBC; very well put together.

**Marxists Philosophy Archive**

[*https://www.marxists.org/reference/subject/philosophy/*](https://www.marxists.org/reference/subject/philosophy/)

Vast collection of out-of-copyright philosophy.

**Berfrois**

[*http://www.berfrois.com/category/main-article/philosophy/*](http://www.berfrois.com/category/main-article/philosophy/)

Berfrois reposts articles from all over the Internet about a variety of subjects relevant to the arts and humanities. It’s hit-and-miss, but well worth a peek from time to time.

**Reddit**

[*https://www.reddit.com/r/philosophy/*](https://www.reddit.com/r/philosophy/)

Interesting articles are posted here from time to time.

## Appendix 2: Reader’s Guide to the Anthology Texts

Below are the anthology texts organised according to the subsections in this resource, which roughly correspond to the specification headings. I have added my own comments on which part of these sometimes long sources I found most useful. I also make tentative recommendations about their accessibility and suitability for use by students as original sources, although I would suggest for all that students have a good grasp of the philosophical arguments and terminology involved before reading.

 Annas, Julia (2006), ‘Virtue Ethics’, in David Copp (ed.), *The Oxford Handbook of Ethical Theory*,Oxford: Oxford University Press, pp. 515–536

 Aristotle, *Nicomachean Ethics*: Books 1, 2, 3, 5, 6, 7, 10 **(Virtue Ethics)**

*Unfortunately, it is thought that as much as two thirds of Aristotle’s works are forever lost to posterity. His entire body of writing is said to have been kept in the Great Library of Alexandria, which is reputed to have been burnt to the ground by Caesar in 40 ce (although this is disputed by modern scholars). What we have left from Aristotle’s oeuvre, which deals as much in biology and rhetoric as it does in ethics and metaphysics, are believed to be his lecture notes. Since they were written for himself and his students, rather than a popular audience (as Plato’s dialogues were), they do not always make for compelling reading. Aristotle could have been one of Ancient Greece’s finest writers, but sadly we will never be able to read him as he intended to be read.*

*Nevertheless, the* Nicomachean Ethics *are a key text for this unit and it is advisable that you at least attempt a reading.*

 Ayer, Alfred J (1973/1991), *The Central Questions of Philosophy*, London, Penguin, pp. 22–29 **(Emotivism)**

 Ayer, Alfred J (1946), *Language, Truth and Logic*, 2nd Edition, New York, Dover (esp. Chapters 1 and 6)

**(Emotivism)**

 Bentham, Jeremy (1879), ‘The Principle of Utility’ in *Introduction to the Principles of Morals and Legislation*, Oxford, Clarendon Press **(Act Utilitarianism)**

*Bentham is fairly straightforward, and, despite being over 200 years old, the language is still easy enough to grasp for a modern reader. It appears as though the board only wants an understanding of the first chapter, which is quite short and readily available online.*

 Diamond, Cora (1978), ‘Eating Meat and Eating People’

 Foot, Philippa (1972), ‘Morality as a system of hypothetical imperatives’ **(Kantian Ethics / Virtue Ethics)**

 Hare, Richard M (1952), *The Language of Morals*, Oxford, Clarendon Press **(Prescriptivism)**

 Hume, David (1739–40), *Treatise of Human Nature*, Book III, Part 1 **(Emotivism)**

*Hume writes well and you are only expected to read a short extract of his work. Worth a look.*

 Kant, Immanuel (1785), *Foundations of the Metaphysics of Morals* **(Kantian Deontological Ethics)**

*Kant’s highly abstract, deeply methodical writing style is infamous among philosophers. He is also fond of inventing new philosophical terms (some of which you have already met; the analytic/synthetic distinction was first explicitly made by Kant). The catchily titled* Foundations of the Metaphysics of Morals *is, however, one of his more accessible works. Deontological ethics forms a large part of this course, so it is worth having at least some familiarity with what Kant actually wrote. For our purposes, the first and second chapters should suffice, although the entire book is not that long.*

 Mackie, J L (1990), ‘The Argument from Queerness in Ethics’ in *Inventing Right and Wrong*, Penguin **(Error Theory)**

 Mill, John Stuart (1863), *Utilitarianism*, Chapters 1, 2, 4, 5

 Moore, George E (1903), *Principia Ethica,* Cambridge University Press,Sections 6–14

 Smart, Jack J C and Williams, Bernard (1973), *Utilitarianism: For and Against*, Chapter 2 (Act

utilitarianism and rule utilitarianism) and Chapter 3 (Hedonistic and non-hedonistic utilitarianism)

The following texts are not specified as set texts in the specification, but will help students to gain deeper understanding of the relevant topics:

 Rachels, J (1993), *The Elements of Moral Philosophy*, 2nd Edition, McGraw-Hill [Particularly Chapter 8, ‘Are There Absolute Moral Rules?’] **(Kantian Deontological Ethics)**

*This book was designed to be an introduction to moral philosophy so it is written in language understandable to the philosophy beginner.*

 Warnock, G J (1967), *Contemporary Moral Philosophy, New Studies in Ethics,* Macmillan – Chapters 1, 3 and 4 **(Intuitionism, Emotivism, Prescriptivism)**

*Sir Geoffrey Warnock’s not so* Contemporary Moral Philosophy *is not particularly easy to get hold of, and, although clear, deals with some fairly advanced concepts. However, if you’re fascinated by meta-ethics it may well be worth tracking down a copy.*

## Appendix 3: A Brief Guide to Writing Philosophy

You are assessed on two assessment objectives on this course:

**AO1**: Demonstrate knowledge and understanding of the core concepts and methods of philosophy, including through the use of philosophical analysis.

**AO2**: Analyse and evaluate philosophical arguments to form reasoned judgements.

**To get AO2 marks…**

• Be critical! Argue! Tell the examiner why you think everyone else is wrong and you are right…

… but always think of *counterarguments* to your position…

… and counterarguments to those counterarguments

***(You want to win the argument after all, right?)***

• Write an introduction…

… and a conclusion.

The introduction should say what you are *going to argue…*

… the conclusion should say what you *have argued.*

***(Don’t change your mind halfway through!)***

**Some general tips…**

• Use *philosophical terms* such as a priori, deductive argument, necessary condition, synthetic and so on…

… but first make sure you understand them! **(*Don’t confuse similar terms!)***

• Worry less about writing *well,* worry more about writing *clearly.*

Many philosophers have been awful writers. The examiner does not want to be moved to tears by the stunning metaphors employed in your essay, they just want to understand it!

• However, it’s still important to have *good grammar.* Philosophical points can completely change in meaning if they are poorly constructed.

If it helps, write your arguments out premise by premise…

I. This argument is written premise by premise

II. Arguments written premise by premise are very clear

III. Therefore, this is a very clear argument

Answers

## Answers to Activities

#### Ethical Theories 1: Utilitarianism

**Dr Achebe’s Fishing Holiday**

Naturally, there isn’t a right or wrong answer here. However, explanations are likely to focus on the various utilitarian justifications for allowing each individual to live / not live. Points which students might raise are the following:

 *Fisherman A is a 57-year-old South American priest who has done much charitable work in his community.*

o If this man is allowed to live, he has a high chance (certainty) of having a positive impact on many people’s lives (extent), who in turn may help others (fecundity)

o On the other hand, he is quite old, so may not have the zeal he once did for charitable work. He is, after all, on a fishing trip (uncertainty, remoteness).

o It may also be argued that the priesthood can be a pernicious as well as a beneficial influence in South America (purity)

 *Fisherman B is a 43-year-old, who works in the accounting department of a large tobacco company. He is happily married with five children.*

o It is likely students will point out the pain tobacco causes throughout the world (extent)

o However, even if the fisherman died his job would probably be filled fairly quickly (uncertainty)

o His wife and his children would also certainly be saddened by his loss; saving his life then is arguably likely to lead to more pleasure on a personal scale than for any of the other candidates (fecundity)

 *Fisherman C is a 27-year-old bachelor who writes screenplays for big-budget Hollywood movies.*

o This fisherman’s films may cause mirth for millions (fecundity)…

o … although the films could be universally panned and hated by audiences (uncertainty)

 *Fisherman D is a 32-year-old widower who works part-time as a handyman. He has one teenage son.*

o This fisherman seems to have very little social impact, and only has one dependent (limited extent and fecundity)…

o … however, the son perhaps only has this one man in the entire world (uncertainty)

o The son will, however, be an adult soon, and we have no indication of what his relationship with his father is like; they may well be estranged or simply not on good terms.

The exercise is intended to demonstrate the limitations of Bentham’s approach; even with four fairly simplistic characters such as these, the number of possibilities which must be taken into account when deciding who lives and who dies quickly grows exponentially upon reflection. A deeper point to take away, along Kantian lines, is that measuring the worth of human life quantitatively is inevitably futile. Instead, we might want to hold, as an absolute principle, that no one life can ever be worth more than any other.

#### Ethical Theories 2: Kantian Deontological Ethics

**Hypothetical and Categorical Imperatives Activity**

Identify whether each of the statements below could be considered a hypothetical imperative or a categorical imperative.

 Put your hands above your head and step away from the vehicle!

o Categorical

 Buy one 12” pizza at Plato’s Pizzeria and get the second absolutely free!

o Hypothetical

 If you want that leather-bound set of Kant’s collected works for Christmas, you’d better finish your philosophy homework.

o Hypothetical

 Everyone should own an Apple product.

o Categorical

 You oughtn’t to wear white trousers if you’re going for a curry this evening.

o Hypothetical

 Listen, you really need to shut up.

o Categorical

 If you love me then you will give me a child.

o Hypothetical

 Worship me!

o Categorical

**First Formulation of the Categorical Imperative Activity**

Identify whether the following activities could be universalised without contradiction:

*NB Many of my answers are open to debate!*

 Stealing bread to feed starving children

o Not universalisable

 Queuing for the cash machine

o Universalisable

 Maintaining a heroin addiction

o Not universalisable

 Having a roast lunch every Sunday

o Universalisable

 Vandalising urinals

o Not universalisable

 Supporting Manchester United

o Not universalisable

 Working as a doctor

o Not universalisable

 Lying about your age and appearance on an online dating site

o Not universalisable

**Second Formulation of the Categorical Imperative Activity**

Identify whether humanity (the agent) is being used as a means or an end in the following scenarios.

 A scout helps a frail old lady across the road to earn his ‘Assisting the Elderly Badge’.

o Means

 A student steals his friend’s assignment so that he can copy his work.

o Means

 On the way home from the pub, Matt gives a homeless man one pound and thirty-three pence.

o Either means or end could be argued

 Belinda compliments her line manager’s terrible outfits and laughs at his bad jokes every day because she wants to be promoted.

o Means

 A woman steals bread from the supermarket to feed her starving family.

o End

 Orhan sleeps with Lucia to make his ex-girlfriend jealous.

o Means

 Steve buys everyone in the office a Milky Bar from the corner shop.

o End

## Answers to Quick Quizzes

#### Ethical Theories 1: Utilitarianism

*1. How does Bentham define utility?*

That which tends to produce pleasure or to reduce pain.

*2. Identify three factors that must be taken into account when calculating pleasure with Bentham’s utility calculus.*

Any three from:

 Its *intensity*

 Its *duration*

 Its *certainty* or *uncertainty*

 Its *propinquity* or *remoteness*

 Its *fecundity*

 Its *purity*

 Its *extent*

*3. How is a higher pleasure to be distinguished from a lower pleasure?*

Pleasure P1 is more desirable than pleasure P2 if: all or almost all people who have had experience of both give a decided preference to P1, irrespective of any feeling that they ought to prefer it.

*4. What is meant by consequentialism?*

Principle that consequences should be the only factor able to determine the moral rightness or wrongness of an action or rule.

*5. Define hedonism.*

Position which holds that the sole good is pleasure.

#### Ethical Theories 2: Kantian Deontological Ethics

*1. What is meant by deontology?*

The study of duty; an ethical theory which holds intentions or unconditional duty are the basis for determining the rightness or wrongness of an action.

*2. Give an example of a hypothetical imperative.*

e.g. If you go to bed early tonight, you can stay up late tomorrow.

*3. What is a maxim?*

A principle or rule to be followed, e.g. ‘Don’t eat too much cake’.

*4. Give an example of a categorical imperative.*

e.g. Stop what you are doing, right now!

*5. State the second formulation of the categorical imperative.*

Act so that persons are always treated as ends, never means.

#### Ethical Theories 3: Aristotle’s Virtue Ethics

*1. Name three of the virtues Aristotle identifies.*

Any three of:

• Courage

• Temperance

• Magnificence

• Magnanimity

• Proper ambition / Pride

• Patience / Good temper

• Truthfulness

• Wittiness

• Friendliness

• Modesty

• Righteous indignation

*2. What word does Aristotle use which is often translated as ‘virtue’?*

*Arête*

*3. What does ‘Eudaimonia’ mean?*

Flourishing, good living, etc.

*4. The golden mean attempts to find a balance between what two things?*

The golden mean holds that a virtue is midway between two vices.

*5. What is ‘practical wisdom’ useful for?*

Chiefly deliberation; in terms of ethics, it involves determining which virtue to exercise in which situation; ‘right time, right place, right action’ and so on.

#### Meta-ethics: Cognitivism and Non-cognitivism

*1. What is cognitivism?*

The philosophical view that moral statements are truth-evaluable.

*2. What is the naturalistic fallacy?*

A criticism of naturalism; that moral properties cannot be reduced to natural properties without being fallacious, or making moral language meaningless.

*3. What is moral anti-realism?*

The belief that there are no mind-independent moral facts or properties.

*4. What is the difference between analytic statements and synthetic statements?*

Analytic statements are tautological, and true simply by the meaning of the terms involved. Synthetic statements are only true by how their meaning corresponds to states of affairs in the world.

*5. What is prescriptivism?*

Prescriptivism is an anti-realist, non-cognitive, meta-ethical view which claims that ethical statements are effectively imperatives that are able to be universalised as judgements.

1. Aristotle, *Nicomachean Ethics*, translated by W D Ross, (The Internet Classics Archive) available at *http://classics.mit.edu//Aristotle/nicomachaen.htm* [↑](#footnote-ref-1)
2. Aristotle, *Nicomachean Ethics*, translated by W D Ross, (The Internet Classics Archive) available at *http://classics.mit.edu//Aristotle/nicomachaen.htm* [↑](#footnote-ref-2)
3. Miller, A, ‘Non-Cognitivism’ in *Routledge Companion to Ethics* ed. by J Skorupski (Oxford: Routledge, 2010). [↑](#footnote-ref-3)
4. Moore, G E, *Principia Ethica* (Cambridge: Cambridge University Press, 1903). [↑](#footnote-ref-4)
5. Hare, R M, *The Language of Morals* (Oxford: Oxford University Press, 1963), p. 127. [↑](#footnote-ref-5)