Session 2

1.7. Fundamental ethical values

1.7.1 Justice

Justice is a concept of moral rightness based ethics, rationality, law, natural law, religion, equity and fairness, as well as the administration of the law, taking into account the inalienable and inborn rights of all human beings and citizens, the right of all people and individuals to equal protection before the law of their civil rights, without discrimination on the basis of race, gender, sexual orientation, gender identity, national origin, color, ethnicity, religion, disability, age, wealth, or other characteristics, and is further regarded as being inclusive of social justice.

According to most contemporary theories of justice, justice is overwhelmingly important: John Rawls claims that "Justice is the first virtue of social institutions, as truth is of systems of thought." Justice can be thought of as distinct from benevolence, charity, prudence, mercy, generosity, or compassion, although these dimensions are regularly understood to also be interlinked. Justice is the concept of cardinal virtues, of which it is one. Justice has traditionally been associated with concepts of fate, reincarnation or Divine Providence, i.e. with a life in accordance with the cosmic plan. The association of justice with fairness has thus been historically and culturally rare and is perhaps chiefly a modern innovation.

Variations of justice

Utilitarianism is a form of consequentialism, where punishment is forward-looking. Justified by the ability to achieve future social benefits resulting in crime reduction, the moral worth of an action is determined by its outcome.

Retributive justice regulates proportionate response to crime proven by lawful evidence, so that punishment is justly imposed and considered as morally correct and fully deserved. The law of

retaliation is a military theory of retributive justice, which says that reciprocity should be equal to the wrong suffered; "life for life, wound for wound, stripe for stripe."

Restorative justice is concerned not so much with retribution and punishment as with:

(a) Making the victim whole

(b) Reintegrating the offender into society.

This approach frequently brings an offender and a victim together, so that the offender can better understand the effect his/her offense had on the victim.

Distributive justice is directed at the proper allocation of things—wealth, power, reward, respect— among different people.

Understandings of justice

Understandings of justice differ in every culture, as cultures are usually dependent upon a shared history, mythology and/or religion. Each culture's ethics create values which influence the notion of justice. Although there can be found some justice principles that are one and the same in all or most of the cultures, these are insufficient to create a unitary justice apprehension.

Justice as harmony

In his dialogue Republic, Plato uses Socrates to argue for justice that covers both the just person and the just City State. Justice is a proper, harmonious relationship between the warring parts of the person or city. Hence Plato's definition of justice is that justice is the having and doing of what is one's own. A just man is a man in just the right place, doing his best and giving the precise equivalent of what he has received. This applies both at the individual level and at the universal level. A person's soul has three parts – reason, spirit and desire. Similarly, a city has three parts – Socrates uses the parable of the chariot to illustrate his point: a chariot works as a whole because the two horses' power

is directed by the charioteer. Lovers of wisdom – philosophers, in one sense of the term – should rule because only they understand what is good. If one is ill, one goes to a doctor rather than a psychologist, because the doctor is expert in the subject of health. Similarly, one should trust one's city to an expert in the subject of the good, not to a mere politician who tries to gain power by giving people what they want, rather than what's good for them. Socrates uses the parable of the ship to illustrate this point: the unjust city is like a ship in open ocean, crewed by a powerful but drunken captain (the common people), a group of untrustworthy advisors who try to manipulate the captain into giving them power over the ship's course (the politicians), and a navigator (the philosopher) who is the only one who knows how to get the ship to port. For Socrates, the only way the ship will reach its destination – the good – is if the navigator takes charge.

Justice as divine command

Advocates of divine command theory argue that justice, and indeed the whole of morality, is the authoritative command of God. Murder is wrong and must be punished, for instance, because, and only because, God commands that it be so.

Divine command theory was famously questioned by Plato in his dialogue, Euthyphro. Called the Euthyphro dilemma, it goes as follows: "Is what is morally good commanded by God because it is morally good, or is it morally good because it is commanded by God?" The implication is that if the former is true, then justice is arbitrary; if the latter is true, then morality exists on a higher order than God, who becomes little more than a passer-on of moral knowledge.

Many apologists have addressed the issue, typically by arguing that is it a false dilemma. For example, some Christian apologists argue that goodness is the very nature of God, and there is necessarily reflected in His commands.[15] Another response, popularized in two contexts by Immanuel Kant and C. S. Lewis, is that it is deductively valid to argue that the existence of an objective morality implies the existence of God and vice versa.

Justice as natural law

For advocates of the theory that justice is part of natural law (e.g., John Locke), it involves the system of consequences that naturally derives from any action or choice. In this, it is similar to the laws of physics: in the same way as the Third of Newton's laws of Motion requires that for every action there must be an equal and opposite reaction, justice requires according individuals or groups what they actually deserve, merit, or are entitled to. Justice, on this account, is a universal and absolute concept: laws, principles, religions, etc., are merely attempts to codify that concept, sometimes with results that entirely contradict the true nature of justice.

Justice as human creation

In contrast to the understandings canvassed so far, justice may be understood as a human creation, rather than a discovery of harmony, divine command, or natural law. This claim can be understood in a number of ways, with the fundamental division being between those who argue that justice is the creation of some humans, and those who argue that it is the creation of all humans.

Justice as trickery In Republic by Plato, the character Thrasymachus argues that justice is the interest of the strong—merely a name for what the powerful or cunning ruler has imposed on the people.

Justice as mutual agreement

According to thinkers in the social contract tradition, justice is derived from the mutual agreement of everyone concerned; or, in many versions, from what they would agree to under hypothetical conditions including equality and absence of bias. This account is considered further below, under 'Justice as fairness'. The absence of bias refers to an equal ground for all people concerned in a disagreement (or trial in some cases).

Justice as a subordinate value

According to utilitarian thinkers including John Stuart Mill, justice is not as fundamental as we often think. Rather, it is derived from the more basic standard of rightness, consequentialism: what is right is what has the best consequences (usually measured by the total or average welfare caused). So, the proper principles of justice are those that tend to have the best consequences. These rules may turn out to be familiar ones such as keeping contracts; but equally, they may not, depending on the facts about real consequences. Either way, what are important are those consequences, and justice is important, if at all, only as derived from that fundamental standard. Mill tries to explain our mistaken belief that justice is overwhelmingly important by arguing that it derives from two natural human tendencies: our desire to retaliate against those who hurt us, and our ability to put ourselves imaginatively in another's place. So, when we see someone harmed, we project ourselves into her situation and feel a desire to retaliate on her behalf. If this process is the source of our feelings about justice that ought to undermine our confidence in them.

1.7.2. Freedom

Freedom often has a different meaning to different people. To a 9 year old, it can be the freedom to choose what to eat or when to brush their teeth.

To a person that lives in a country which is controlled by an oppressive regime, it can mean not living in fear. Freedom is sought after by so many people, yet so few of those people agree on what defines freedom.

What is the fascination with freedom?

As a result of the definition of freedom being different to so many people and because there are so many concepts of freedom, it is always something that seems impossible to achieve.

Human nature has proven that people always want what they cannot have. A poor person is not free to do as they choose because they have no money.

A rich and famous person is not free to do as they choose because of the negative press they may attract. There are always constraints placed on people's actions. This can be because of a person's standards, but more often than not it is because of their peers' perceptions.

Conflicting Freedom Definitions

There are government departments in most countries in the world that are dedicated to enabling and protecting the freedom of their citizens. These government agencies are financed by people who pay taxes, which could be considered contradictory to many freedom definitions.

Taxes are a considerable portion of why people feel a lack of freedom. People who look after their finances and save money in banks will get taxed on the interest on their savings. People who save in offshore accounts will often pay no tax the local authority of the offshore tax haven.

The same can be said of taxes on profits received thanks to investments in precious metals or gems.

Most countries have freedom of speech which allows people to voice their opinions. People can speak out about anything they disagree with or even instigate campaigns to protest against something they see as an injustice. This is usually a fair and civilized way of life, but the freedom to speak out can sometimes have a negative impact on others.

This is when the greater good has to be assessed and the decision to restrict what some people can express is the right thing to do. In this situation the definition of freedom will not be the same as the definition of freedom in the everyday sense. In this scenario, people are free to do what they want as long as what they do does not affect others.

Liberty and the Justice System

People have been policed and required to abide by the laws of their country or state for hundreds of years. These laws mean that people are not free to act in ways that the general public consider detrimental to others. When people break these laws, they risk having their freedom or liberty taken away for a period of time. It is only when some people are incarcerated that they first think about freedom.

Achieving Freedom

The search for freedom can be a complex task that requires a person to consider the freedom definition in all of the above scenarios. A person can then accurately assess what the freedom definition would be for their own life.

By really understanding what freedoms a person already has compared to people in other situations, the feeling of freedom can be a lot easier to grasp than they originally thought.

1.7.3. Trust

The commonly understood opposite of truth is falsehood, which, correspondingly, can also take on a logical, factual, or ethical meaning. The concept of truth is discussed and debated in several contexts, including philosophy and religion. Many human activities depend upon the concept, which is assumed rather than a subject of discussion, including science, law, and everyday life.

Various theories and views of truth continue to be debated among scholars, philosophers, and theologians. Language and words are a means by which humans convey information to one another and the method used to determine what is a "truth" is termed a criterion of truth. There are differing claims on such questions as what constitutes truth: what things are truth bearers capable of being true or false; how to define and identify truth; the roles that faith-based and empirically based knowledge play; and whether truth is subjective or objective, relative or absolute.

Major theories of truth

The question of what is a proper basis for deciding how words, symbols, ideas and beliefs may properly be considered true, whether by a single person or an entire society, is dealt with by the five most prevalent substantive theories listed below. Each theory presents perspectives that are widely shared by published scholars.

However, the substantive theories are not universally accepted. More recently developed "deflationary" or "minimalist" theories of truth have emerged as competitors to the older substantive theories. Minimalist reasoning centers on the notion that the application of a term like true to a statement does not assert anything significant about it, for instance, anything about its nature. Minimalist reasoning realizes truth as a label utilized in general discourse to express agreement, to stress claims, or to form general assumptions.

Substantive theories

Correspondence theories emphasize that true beliefs and true statements correspond to the actual state of affairs. This type of theory stresses a relationship between thoughts or statements on one hand, and things or objects on the other. It is a traditional model tracing its origins to ancient Greek philosophers such as Socrates, Plato, and Aristotle. This class of theories holds that the truth or the falsity of a representation is determined in principle entirely by how it relates to "things", by whether it accurately describes those "things." An example of correspondence theory is the statement by the Thirteenth Century philosopher/theologian Thomas Aquinas: Veritas est adaequatio rei et intellectus ("Truth is the equation or adequation of things and intellect"), a statement which Aquinas attributed to the Ninth Century neoplatonist Isaac Israeli. Aquinas also restated the theory as: "A judgment is said to be true when it conforms to the external reality"

Correspondence theory centers heavily around the assumption that truth is a matter of accurately copying what is known as "objective reality" and then representing it in thoughts, words and other symbols. Many modern theorists have stated that this ideal cannot be achieved without analyzing

additional factors. For example, language plays a role in that all languages have words to represent concepts that are virtually undefined in other languages. The German word Zeitgeist is one such example: one who speaks or understands the language may "know" what it means, but any translation of the word apparently fails to accurately capture its full meaning (this is a problem with many abstract words, especially those derived in agglutinative languages). Thus, some words add an additional parameter to the construction of an accurate truth predicate. Among the philosophers who grappled with this problem is Alfred Tarski, whose semantic theory is summarized further below in this article.

Proponents of several of the theories below have gone further to assert that there are yet other issues necessary to the analysis, such as interpersonal power struggles, community interactions, personal biases and other factors involved in deciding what is seen as truth.

Coherence theory

Walter Seymour All ward's Veritas (Truth) outside Supreme Court of Canada, Ottawa, Ontario Canada For coherence theories in general, truth requires a proper fit of elements within a whole system. Very often, though, coherence is taken to imply something more than simple logical consistency; often there is a demand that the propositions in a coherent system lend mutual inferential support to each other. So, for example, the completeness and comprehensiveness of the underlying set of concepts is a critical factor in judging the validity and usefulness of a coherent system. A pervasive tenet of coherence theories is the idea that truth is primarily a property of whole systems of propositions, and can be ascribed to individual propositions only according to their coherence with the whole. Among the assortment of perspectives commonly regarded as coherence theory, theorists differ on the question of whether coherence entails many possible true systems of thought or only a single absolute system.

Some variants of coherence theory are claimed to describe the essential and intrinsic properties of formal systems in logic and mathematics. However, formal reasons are content to

contemplate axiomatically independent and sometimes mutually contradictory systems side by side, for example, the various alternative geometries. On the whole, coherence theories have been rejected for lacking justification in their application to other areas of truth, especially with respect to assertions about the natural world, empirical data in general, assertions about practical matters of psychology and society, especially when used without support from the other major theories of truth.

Coherence theories distinguish the thought of rationalist philosophers, particularly of Spinoza, Leibniz, and G.W.F. Hegel, along with the British philosopher F.H. Bradley. They have found a resurgence also among several proponents of logical positivism, notably Otto Neurath and Carl Hempel.

Constructivist theory

Social constructivism holds that truth is constructed by social processes, is historically and culturally specific, and that it is in part shaped through the power struggles within a community. Constructivism views all of our knowledge as "constructed," because it does not reflect any external "transcendent" realities (as a pure correspondence theory might hold). Rather, perceptions of truth are viewed as contingent on convention, human perception, and social experience. It is believed by constructivists that representations of physical and biological reality, including race, sexuality, and gender, are socially constructed.

Giambattista Vico was among the first to claim that history and culture were man-made. Vico's epistemological orientation gathers the most diverse rays and unfolds in one axiom – verum ipsum factum – "truth itself is constructed". Hegel and Marx were among the other early proponents of the premise that truth is, or can be, socially constructed. Marx, like many critical theorists who followed, did not reject the existence of objective truth but rather distinguished between true knowledge and knowledge that has been distorted through power or ideology. For Marx, scientific and true knowledge

is "in accordance with the dialectical understanding of history" and ideological knowledge is "an epiphenomenal expression of the relation of material forces in a given economic arrangement".

Consensus theory

Consensus theory holds that truth is whatever is agreed upon, or in some versions, might come to be agreed upon, by some specified group. Such a group might include all human beings, or a subset thereof consisting of more than one person.

Among the current advocates of consensus theory as a useful accounting of the concept of "truth" is the philosopher Jürgen Habermas. Habermas maintains that truth is what would be agreed upon in an ideal speech situation. Among the current strong critics of consensus theory is the philosopher Nicholas Rescher.

Pragmatic theory

The three most influential forms of the pragmatic theory of truth were introduced around the turn of the 20th century by Charles Sanders Peirce, William James, and John Dewey. Although there are wide differences in viewpoint among these and other proponents of pragmatic theory, they hold in common that truth is verified and confirmed by the results of putting one's concepts into practice.

Peirce defines truth as follows: "Truth is that concordance of an abstract statement with the ideal limit towards which endless investigation would tend to bring scientific belief, which concordance the abstract statement may possess by virtue of the confession of its inaccuracy and one-sidedness, and this confession is an essential ingredient of truth." This statement stresses Peirce's view that ideas of approximation, incompleteness, and partiality, what he describes elsewhere as fallibilism and "reference to the future", are essential to a proper conception of truth. Although Peirce uses words like concordance and correspondence to describe one aspect of the pragmatic sign relation, he is also quite

explicit in saying that definitions of truth based on mere correspondence are no more than nominal definitions, which he accords a lower status than real definitions.

William James's version of pragmatic theory, while complex, is often summarized by his statement that "the 'true' is only the expedient in our way of thinking, just as the 'right' is only the expedient in our way of behaving." By this, James meant that truth is a quality, the value of which is confirmed by its effectiveness when applying concepts to practice (thus, "pragmatic").

John Dewey, less broadly than James but more broadly than Peirce, held that inquiry, whether scientific, technical, sociological, philosophical or cultural, is self-corrective over time if openly submitted for testing by a community of inquirers in order to clarify, justify, refine and/or refute proposed truths.

Though not widely known, a new variation of the pragmatic theory was defined and wielded successfully from the 20th century forward. Defined and named by William Ernest Hocking, this variation is known as "negative pragmatism". Essentially, what works may or may not be true, but what fails cannot be true because the truth always works. Richard Feynman also ascribed to it: "We never are definitely right, we can only be sure we are wrong." This approach incorporates many of the ideas from Peirce, James, and Dewey. For Peirce, the idea of "... endless investigation would tend to bring about scientific belief ..." fits negative pragmatism in that a negative pragmatist would never stop testing. As Feynman noted, an idea or theory "... could never be proved right, because tomorrow's experiment might succeed in proving wrong what you thought was right." Similarly, James and Dewey's ideas also ascribe truth to repeated testing which is "self-corrective" over time.

Pragmatism and negative pragmatism are also closely aligned with the coherence theory of truth in that any testing should not be isolated but rather incorporate knowledge from all human endeavors and experience. The universe is a whole and integrated system, and testing should acknowledge and account for its diversity. As Feynman said, "... if it disagrees with experiment, it is wrong."

Minimalist

Modern developments in the field of philosophy, starting with the relatively modern notion that a theory being old does not necessarily imply that it is completely flawless, have resulted in the rise of a new thesis: that the term truth does not denote a real property of sentences or propositions. This thesis is in part a response to the common use of truth predicates (e.g., that some particular thing "...is true") which was particularly prevalent in philosophical discourse on truth in the first half of the 20th century. From this point of view, to assert that "'2 + 2 = 4' is true" is logically equivalent to asserting that "2 + 2 = 4", and the phrase "is true" is completely dispensable in this and every other context. In common parlance, truth predicates are not commonly heard, and it would be interpreted as an unusual occurrence was someone to utilize a truth predicate in an everyday conversation when asserting that something is true. Newer perspectives that take this discrepancy into account and work with sentence structures that are actually employed in common discourse can be broadly described:

As deflationary theories of truth, since they attempt to deflate the presumed importance of the words "true" or truth, as disquisitional theories, to draw attention to the disappearance of the quotation marks in cases like the above example, or as minimalist theories of truth.

Whichever term is used, deflationary theories can be said to hold in common that "he predicate 'true' is an expressive convenience, not the name of a property requiring deep analysis."

Once we have identified the truth predicate's formal features and utility, deflationists argue, we have said all there is to be said about truth. Among the theoretical concerns of these views is to explain away those special cases where it does appear that the concept of truth has peculiar and interesting properties.

In addition to highlighting such formal aspects of the predicate "is true", some deflationists point out that the concept enables us to express things that might otherwise require infinitely long sentences.

1.7.4. Responsibility

We evaluate people and groups as responsible or not, depending on how seriously they take their responsibilities. Often we do this informally, via moral judgment. Sometimes we do this formally, for instance in legal judgment. This article considers mainly moral responsibility, and focuses largely upon individuals. Later sections also comment on the relation between legal and moral responsibility, and on the responsibility of collectives.

The article discusses four different areas of individual moral responsibility:

(1) Responsible agency, whereby a person is regarded as a normal moral agent;

(2) Retrospective responsibility, when a person is judged for her actions, for instance, in being blamed or punished;

(3) Prospective responsibility, for instance, the responsibilities attaching to a particular role; and

(4) Responsibility as a virtue, when we praise a person as being responsible.

Philosophical discussion of responsibility has focused largely on (1) and (2). The article points out that a wider view of responsibility helps explore some connections between moral and legal responsibility, and between individual and collective responsibility. It also enables us to relate responsibility to its original philosophical use, which was in political thought.

1. Introduction

The original philosophical usage of "responsibility" was political (see McKeon, 1957). This reflected the origin of the word. In all modern European languages, "responsibility" only finds a home toward the end of the eighteenth century. This is within debates about representative government, that is, government which is responsible to the people. In the etymology of "responsibility," the Oxford English Dictionary cites the debates on the U.S. constitution in the Federalist Papers (1787), and the

Anglo-Irish political thinker Edmund Burke (1796). When John Stuart Mill writes of responsibility, in the middle of the nineteenth century, again his concern is not with free will, but with the principles of representative government. At the end of the nineteenth century, the most notable thinker to speak of responsibility is Max Weber, who propounds an ethics of responsibility (Verantwortungsethik) for the politician. For Weber, the vocation of politics demands a calm attention to the facts of the situation and the consequences of actions – and not to lofty or abstract principles.

So far as responsibility has a place in eighteenth and nineteenth century thought, then, this is in political contexts, where the concern is with responsible action and the principles of representative government. In twentieth century philosophy, on the other hand, the emphasis has been on questions of free will and determinism: Is a person responsible for her actions or character? Would the truth of determinism eliminate such responsibility? Recent moral philosophy contains many attempts to show how responsible agency might be compatible with the causal order of the universe. These debates obviously center on the individual agent. As such, they pose difficulties for understanding the topic of collective responsibility – an issue that twentieth century politics has raised with a new urgency. Nor does a concern with free will correspond to many everyday issues about responsibility – for example, questions of mutual accountability, defining a person's sphere of responsibility, or judging a person as sufficiently responsible for a particular role.

In fact, there are several important uses of responsibility as it relates to individuals, which this article will tackle in turn. There are also important questions about the distinction between moral and legal responsibility. The article will then consider what relations there may be between the concept's individual and collective uses. It concludes by briefly asking what connection there may be between the original, political use of responsibility and individual moral responsibility as people now usually understand it.

2.   **Individual Responsibility**

There is no philosophically well-settled way of dividing or analyzing the various components of responsibility, and some components are often ignored by philosophers. To take a more comprehensive approach, this article divides the responsibility of individuals into four areas of enquiry. Recent analytic moral philosophy has tended to ask two deceptively simple questions about responsibility:

"What is it to be responsible?" and

"What is a person responsible for?"

The first question is usually taken as a question about moral agency, the second as a question about holding people accountable for past actions. As noted, however, this does not capture the variety of uses that we make of the concept. We can see this by observing that both questions might mean something quite different, leading us to four distinct topics, as follows:

"What is it to be responsible?" is most often asked by philosophers as a question about the foundations of moral agency. What sort of creature can properly be held responsible for its actions? The simple answer is: a normal human adult. To explain and justify this reply, philosophers tend to turn to psychological and metaphysical features of normal adults, such as free will. We might also approach the same issue with a somewhat different emphasis: What features of (normal, adult) human interaction are involved in our holding one another responsible?

However, in asking "What is it to be responsible?" we might also have a second question in mind. We often praise some people as responsible, and criticize others as irresponsible. Here responsibility names a virtue – a morally valuable character trait. We may also praise an institution as responsible. One of the word's original uses was to call for "responsible government." We can compare this with the more recent demand that corporations be "socially responsible." This aspect of responsibility has received very little philosophical attention.

"What is a person responsible for?" is a question most often asked by philosophers in connection with causation and accountability. This retrospective, or backward-looking, use is closely connected with praise and blame, punishment, and desert. When something has gone wrong, we invariably want to know who was at fault; and when something has gone right, we occasionally stop to ask who acted well. This is the topic of retrospective responsibility.

Again, however, we might use the same words to ask an entirely different question: "What is a person responsible for?" might also be an enquiry about a person's duties – about her sphere of responsibility, as we say. A parent is responsible for caring for his child, an employee for doing her job, a citizen for obeying the law. It is a basic fact of human cooperation that responsibilities are often divided up between people: for example, the doctor is responsible for prescribing the right drugs, and the patient responsible for taking them correctly. As against questions of retrospective responsibility, this topic is sometimes termed prospective responsibility, that is, what responsibilities we are duty-bound to undertake.

These two apparently simple questions ("What is it to be responsible?" and "What is a person responsible for?") about individual responsibility thus point to four different topics:

Moral agency

Responsibility as a virtue

Retrospective responsibility

Prospective responsibility

Each of these topics poses a host of important philosophical questions. Both the retrospective and prospective uses also raise the relation between legal and moral responsibility. Many important

theories of responsibility relate to legal concerns, which will be discussed in a later section. As we pursue these topics, there is also the difficulty of seeing how they interrelate, so that it makes sense that we use the same word to raise each issue.

The discussion begins with the topics which philosophers have most often discussed: the nature of moral agency and retrospective responsibility.

A. Moral Agency

Normal human adults represent our paradigm case of responsible agents. What is distinctive about them, that we accord them this status? Thinking of retrospective responsibility in particular, why can be held accountable for their actions – justly praised or blamed, deservedly punished or rewarded? The philosophical literature has explored three broad approaches to moral agency:

Human beings have free will, that is, distinctive causal powers or a special metaphysical status, that separate them from everything else in the universe;

Human beings can act on the basis of reason(s);

Human beings have a certain set of moral or proto-moral feelings.

The first approach, although historically important, has largely been discredited by the success of modern science. Science provides, or promises, naturalistic explanations of such phenomena as the evolution of the human species and the workings of the brain. Almost all modern philosophers approach responsibility as compatibilists – that is, they assume that moral responsibility must be compatible with causal or naturalistic explanation of human thought and action, and therefore reject the metaphysical idea of free will. (An important note: There can be terminological confusion here. Some contemporary philosophers will use the term "free will" to describe our everyday freedom of choice, claiming that free will, properly understood, is compatible with the world's causal order.)

Among modern compatibilists, a contest remains, however, between the second and third approaches – positions that are essentially Kantian and Humean in inspiration. Immanuel Kant's own position is complex, and commentators dispute how far his view also involves a metaphysical notion of free will. It is indisputable, however, that our rationality is at the Centre of his picture of moral agency. Kant himself does not speak of responsibility – the word was only just coming into the language of his day – but he does have much to say about imputation (Zurechnung), that is, the basis on which actions are imputed to a person. Kant was principally concerned with evaluation of the self. Although he occasionally mentions blame (mutual accountability), his moral theory is really about the basis on which a person treats herself as responsible. The core of his answer is that a rational agent chooses to act in the light of principles – that is, we deliberate among reasons. Therefore standards of rationality apply to us, and when we fail to act rationally this is, simply and crudely, a Bad Thing. It is important to be aware that Kant sees reason as having moral content, so that there is a failure of rationality involved when we do something immoral – for instance, by pursuing our self-interest at the expense of others. Even if we sometimes feel no inclination to take account of others, reason still tells us that we should, and can motivate us to do so. Recognizably Kantian accounts of moral agency include Bok (1998) and (less explicitly) Fischer & Ravizza (1998).

The issue of reason's moral content separates Kantians from Humeans. David Hume denied that reason can provide us with moral guidance, or the motivation to act morally. He is famous for his claim that "Reason is wholly inactive, and can never be the source of so active a principle as conscience, or a sense of morals" (A Treatise of Human Nature, book 3, part 1, sect. 1). If we are moral agents, this is because we are equipped with certain tendencies to feel or desire, dispositions that make it seem rational to us to act and think morally. Hume himself stressed our tendency to feel sympathy for others and our tendency to approve of actions that lead to social benefits (and to disapprove of those contrary to the social good). Another important class of feelings concerns our tendencies to feel shame or guilt, or more broadly, to be concerned with how others see our actions

and character. A Humean analysis of responsibility will investigate how these emotions lead us to be responsive to one another, in ways that support moral conduct and provide social penalties for immoral conduct. That is, its emphasis is less on people's evaluation of themselves and more on how people judge and influence one another. Russell (1995) carefully develops Hume's own account. In twentieth century philosophy, broadly Humean approaches were given a new lease of life by Peter Strawson's "Freedom and Resentment" (1962). This classic essay underlined the role of "reactive sentiments" or "reactive attitudes" – that is, emotional responses such as resentment or shame – in practices of responsibility.

The basic criticisms that each position makes of the other are simple. Kantians are vulnerable to the charge that they do not give a proper account of the role of feeling and emotion in the moral life. They can also be accused of reifying our capacity for reason in a way that makes mysterious how human beings' capacities for reason and morality might have evolved. Humeans are vulnerable to the charge that they cannot give any account of the validity of reasoning beyond the boundaries of what we might feel inclined to endorse or reject: Can the Humean really hold that moral reasoning has any validity for people who do not feel concern for others? Contemporary philosophers have developed both positions so as to take account of such criticisms, which has led to rather technical debates about the nature of reason (for instance, Bernard Williams' (1981) well-known distinction between internal and external reasons) and normativity (what it is for something to provide a reason to act or think in a certain way, for example, Korsgaard, 1996). So far as responsibility is concerned, Wallace (1994) is a well-regarded attempt to mediate between the two approaches. Rather differently, Pettit (2001) uses our susceptibility to reasons as the basis for an essentially interactive account of moral agency.

For our purposes, perhaps the most important point is that both positions highlight a series of factors important to responsibility and mutual accountability. These factors include: general responsiveness to others (for instance, via moral reasoning or feelings such as sympathy); a sense of responsibility for our actions (for instance, so that we may offer reasons for our actions or feel emotions

of shame or guilt); and tendencies to regard others as responsible (for instance, to respect persons as the authors of their deeds and to feel resentful or grateful to them). In each case, note that the first example in brackets has a typically Kantian (reason-based) cast, the second a Humean (feeling/emotion-related) cast.

Two further thoughts should be added which apply regardless of which side of this debate one inclines toward. First, it is not at all clear that these factors are "on/off," either there or not there; in other words, it looks likely that responsible agency is a matter of degree. One possible implication of this is that some other animals might have a degree of moral agency; another implication is that human beings may vary in the extent of their agency. (This seems clearly true of children as opposed to adults.

We may be more reluctant to believe that the extent of adults' moral agency can vary, but such a claim is not obviously false.) Second, none of these factors has an obvious connection to free will, in the metaphysical sense that opposes free will to determinism. As we shall see, however, whether we emphasize the rational or the affective basis for responsible agency tends to generate characteristically different accounts of retrospective responsibility, where the issue of free will tends to recur.

b. Retrospective Responsibility

In assigning responsibility for an outcome or event, we may simply be telling a causal story. This might or might not involve human actions. For example: the faulty gasket was responsible for the car breaking down; his epileptic fit was responsible for the accident. Such usages do not imply any assignment of blame or desert, and philosophers often distinguish them by referring to "causal responsibility." More commonly, however, responsibility attribution is concerned with the morality of somebody's action(s). Among the many different causes that led to an outcome, that action is identified as the morally salient one. If we say the captain was responsible for the shipwreck, we do not deny that all sorts of other causes were in play. But we do single out the person who we think ought to be held responsible for the outcome. Philosophers sometimes distinguish this usage, by speaking of "liability responsibility."

Retrospective responsibility usually involves, then, a moral (or perhaps legal) judgment of the person responsible. This judgment typically pictures the person as liable to various consequences: to feeling remorse (or pride), to being blamed (or praised), to making amends (or receiving gratitude), and so forth.

This topic is an old concern of philosophers, predating the term "responsibility" by at least two millennia. The classic analysis of the issues goes back to Aristotle in the Nicomachean Ethics, where he investigates the conditions that exculpate us from blame and the circumstances where blame is appropriate. Among conditions that excuse the actor, he mentions intoxication, force of circumstances, and coercion: we cannot be held responsible where our capacity to choose was grossly impaired or where there was no effective choice open to us (though perhaps we can be blamed for getting into that condition or those circumstances). We can be blamed for what we do when threatened by others, but not as we would be if coercion were absent. In each case, the issue seems to be whether or not we are able to control what we do: if something lies beyond our control, it also lies beyond the scope of our responsibility.

However, although Aristotle thinks that our capacities for deliberation and choice are important to responsible agency, he lacks the Kantian emphasis on rational control discussed in the last section. Aristotle grants considerable importance to habituation and stable character traits – the virtues and vices. Hence another way of interpreting what he says about responsibility is to argue that Aristotle's excusing conditions represent cases where an action does not reveal a person's character: everybody would act like that if circumstances provided no other choice; no one makes responsible choices when drunk. On the other hand, how we respond to coercion does reveal much about our virtues and vices; the point is that the meaning of such acts is very different from the meaning they would have in the absence of coercion.

In its emphasis on character, Aristotle's account is much closer to Hume's than to Kant's, since character is about tendencies to feel and behave in various ways, as well as to think and choose. Given that Kant's moral psychology is usually thought to be less plausible than Aristotle or Hume's, it is interesting that Kantian approaches have, nonetheless, dominated modern approaches to retrospective responsibility. Why should this be so?

Kant's underlying thought is that the person who acts well deserves to be happy (he continually refers to goodness as "worthiness to be happy"). The person who acts badly does not: she deserves to be reproached, ought to feel remorse, and may even deserve punishment. Since blame, guilt and punishment are of great practical importance, it is clearly desirable that our account of responsibility justify them. Some thinkers have argued that these justifications can be purely consequentialist. For instance, Smart (1961) argues that blame, guilt and punishment are only merited insofar as they can encourage people to do better in the future. However, most philosophers have been dissatisfied with such accounts. Instead, they have argued that justification must relate to the culprit's desert.

For most people, the intuitive justification for the sort of desert involved in retrospective responsibility lies in individual choice or control. You chose to act selfishly: you deserve blame. You chose not to take precautions: you deserve to bear the consequences. You chose to break the law: you deserve punishment. (The question of legal responsibility is considered separately, below.) This way of putting matters clearly gives pride of place to our capacity to control our conduct in the light of reasons, moral and otherwise. It will also emphasize the intentions underlying an action rather than its actual outcomes. This is because intentions are subject to rational choice in a way that outcomes often are not. Kant's thought that the rational agent can choose whether or not to act on the basis of reasons is sometimes expressed in the idea that we should each be respected as the authors of our thoughts and intentions. This thought has the less positive consequence that when somebody chooses immorally and irrationally, he fails in a distinctive way, so that he is not (in Kant's terms) "worthy to be happy." Note, however, that this line of thought is open to a very obvious objection. It can be argued that our

intentions and choices are conditioned by our characters, and our characters by the circumstances of our upbringing. Clearly these are not matters of choice. This is why a concern with retrospective responsibility raises the family of issues around moral luck and continues to lead back to the issue of free will: the idea that we are, really and ultimately, the authors of our own choices – despite scientific and common-sense appearances.

Contrasting Kant's approach with that of Aristotle and utilitarianism. Humeans, favoring naturalistic explanation of thought and action, are likely to be drawn to elements of the last two – namely Aristotle's emphasis on actions as revealing virtues and vices, and the consequentialist emphasis on social benefits of practices of accountability. In particular, Humeans are much more likely to see retrospective responsibility in terms of the feelings that are appropriate – for instance, our resentment at someone's bad conduct, or our susceptibility to shame at others' responses. Clearly, such feelings and the resulting actions are about our exercising mutual influence on one another's conduct for the sake of more beneficial social interaction. In other words, although the Humean analysis can be understood in terms of individual psychology, it also points to the question: What is it about human interaction that leads us to hold one another responsible? Kantians, on the other hand, tend to think of retrospective responsibility, not as a matter of influencing others, but rather as our respecting individual capacities for rational choice. This respect may still have harsh consequences, as it involves granting people their just deserts, including blame and punishment.

c. Prospective Responsibility

A different use of "responsibility" is as a synonym for "duty." When we ask about a person's responsibilities, we are concerned with what she ought to be doing or attending to. Sometimes we use the term to describe duties that everyone has – for example, "Everyone is responsible for looking after his own health." More typically, we use the term to describe a particular person's duties. He is responsible for sorting the garbage; she is responsible for looking after her baby; the Environmental

Protection Agency is responsible for monitoring air pollution; and so on. In these cases, the term singles out the duties, or "area of responsibility," that somebody has by virtue of their role.

This usage bears at least one straightforward relation to the question of retrospective responsibility. We will tend to hold someone responsible when she fails to perform her duties. A captain is responsible for the safety of the ship; hence he will be held responsible if there is a shipwreck. The usual justification for this lies in the thought that if he had taken his responsibility more seriously, then his actions might have averted the shipwreck. In some cases, though, when we are entrusted with responsibility for something, we will be held responsible if harm occurs, regardless of whether we might have averted it. This might be true if one hires (that is, rents) a car, for instance: even if an accident is not your fault, the contract may stipulate that you will be responsible for part of the repair costs. In order to hire (rent) the car in the first place, one must accept – take responsibility for – certain risks.

Legal thinkers, in particular, have pointed out that this suggests one way in which Kantian approaches – that is, approaches to responsibility which focus on acts and outcomes that were under a person's control – may be inadequate. We may think that everybody has a duty (that is, a prospective responsibility) to make recompense when certain sorts of risks materialize from their actions. Consider a standard example: suppose John accidentally slips and breaks a vase in Jane's shop. This is probably not something John had control over, and to avoid the risk of damaging any of Jane's possessions, John would have to avoid entering her shop altogether. Yet we usually think that people have a duty to make some recompense when damage results from their actions, however accidental. From the point of view of our interacting with one another, the issue is not really whether a person could have avoided a particular, unfortunate outcome, so much as the fact that all our actions create risks; and when those risks materialize, someone suffers. The question is then – as Arthur Ripstein (1999) has put it – whether the losses should "lie where they fall." To say that they should is basically to shrug our shoulders about the damage; in that case, the only person who suffers is the shop-owner. But we

often think that losses should be redistributed. For that to happen someone else has to make some sort of amends – in this case, the person who caused the accident will have to accept responsibility.

In terms of prospective responsibility, then, we may think that everyone has a duty to make certain amends when certain risks of action actually materialize – just because all our actions impose risks on others as well as ourselves. In this case, retrospective responsibility is justified, not by whether the person controlled the outcome or could have chosen to do otherwise, but by reference to these prospective responsibilities. Notice, however, that we might want to distinguish the duty to make amends from the issue of blameworthiness. One might accept the above account as to why the customer should compensate the owner of the broken vase, but add that in such a case she is not to blame for the breakage. There is clearly some merit to this response. It suggests that retrospective responsibility is more complicated than is often thought: blameworthiness and liability to compensate are different things, and may need to be justified in different ways. However, this question has not really been systematically pursued by moral philosophers, although the distinction between moral culpability and liability to punishment has attracted much attention among legal philosophers.

The connection between prospective and retrospective responsibility raises another complication. This stems from the fact that people often disagree about what they ought to do – that is, about what people's prospective responsibilities are. This question of moral disagreement is not often mentioned in debates about responsibility, but may be rather important. To take an example: people have very different beliefs about the ethics of voluntary euthanasia – some call it "mercy killing," others outright murder. Depending on our view, we will tend to blame or to condone the person who kills to end grave suffering. In other words, different views of somebody's prospective responsibilities will lead to very different views of how retrospective responsibility ought to be assigned. One might even argue that many of our moral disagreements are actually brought to light, and fought out, when actors and on-lookers dispute what responses are appropriate. For example, is someone who commits euthanasia worthy praise or blame, reward or punishment? These disagreements, often very vocal, are important

for the whole topic of responsibility, because they relate to how moral agents come to be aware of what morality demands of them.

Kantian ethics typically describes moral agency in terms of the co-authorship of moral norms: the rational agent imposes norms upon her, and so can regard herself as an "author" of morality. (This element of Kantian ethics can be difficult to appreciate, because Kant is so clear that everyone should impose the same objective morality on themselves.) Whether or not one accepts the Kantian emphasis upon rationality or a Universalist morality, it is clear that an important element of responsible agency consists in judging one's own responsibilities. Hence, we do not tend to describe a dutiful child as responsible. This is because he obeys, rather than exercising his own judgment about what he ought to do. This issue is not just about how we judge our own duties, however: it's also about how others judge us, and our right to judge others. So far as others regard us as responsible, they will recognize that we also have a right to judge what people's prospective responsibilities are, and how retrospective responsibility ought to be assigned. Importantly, people can recognize one another as responsible in this way, even in the face of quite deep moral disagreements. By the same token, we know how disrespectful it is of someone, not to take her moral judgments seriously.

The question of how far we are entitled to judge prospective responsibilities – our own and other people's – and how far we are entitled to judge retrospective responsibilities – our own and others' – raises yet another complication for how we think about responsibility. As the example of childhood suggests, there can be degrees of responsibility. Ascribing different degrees of responsibility may be necessary or appropriate with regard to different sorts of decision-making. Hence we sometimes say, "He's not ready for that sort of responsibility" or "She couldn't be expected to understand the implications of that sort of choice." In the first place, such statements highlight the close connection between prospective and retrospective responsibility: it will not be appropriate to hold someone (fully) responsible for his actions if he was faced with responsibilities that were unrealistic and over-demanding. It also points to the fact that people vary in their capacities to act and judge

responsibility. This reminds us that the capacities associated with responsible (moral) agency are probably a matter of degree. It might also remind us of a fourth use of "responsibility": to name a virtue of character.

d. Responsibility as a Virtue

While theories of moral agency tend to regard an agent as either responsible or not, with no half-measures, our everyday language usually deploys the term "responsible" in a more nuanced way. As just indicated, one way we do this is by weighing degrees of responsibility, both with regard to the sort of prospective responsibilities a person should bear and a person's liability to blame or penalties. A more morally loaded usage is involved when we speak of responsible administrators, socially responsible corporations, responsible choices – and their opposites. In these cases, we use the term "responsible" as a term of praise: responsibility represents a virtue that people (and organizations) may exhibit in one area of their conduct, or perhaps exemplify in their entire lives.

In such cases, our meaning is usually quite clear. The responsible person can be relied on to judge and to act in certain morally desirable ways; in the case of more demanding ("more responsible") roles, the person can be trusted to exercise initiative and to demonstrate commitment; and when things go wrong, such a person will be prepared to take responsibility for dealing with things. One way of putting this might be to say that the responsible person can be counted on take her responsibilities seriously. We will not need to hold her responsible, because we can depend on her holding herself responsible. Another way of putting the matter would be much more contentious, and harkens back to the question of whether we should think of moral agency as a matter of degree. One might claim that the responsible person possesses the elements pertaining to moral agency (such as capacities to judge moral norms or to respond to others) to a greater degree than the irresponsible person. This would be highly controversial, because it seems to undermine the idea that all human beings are equal moral

agents. However, it would help us to see why a term we sometimes use to describe all moral agents can also be used to praise some people rather than others.

However this may be, it is fair to say that this usage of "responsible" has received the least attention from philosophers. This is interesting given that this is clearly a virtue of considerable importance in modern societies. At any rate, it is possible to see some important connections between the virtue and the areas that philosophers have emphasized.

The irresponsible person is not one who lacks prospective responsibilities, nor is she one who may not be held responsible retrospectively. It is only that she does not take her responsibilities seriously. Note, however, that the more responsible someone is, the more we will be inclined to entrust her with demanding roles and responsibilities. In this case, her "exposure," as it was, to being held retrospectively responsible increases accordingly. And the same is true in the opposite direction, when someone consistently behaves less responsibly. An illuminating essay by Herbert Fingarette (1967) considers the limit case of the psychopath, someone who shows absolutely no moral concern for others, nor any sensitivity to moral reproach. Perhaps our first response will be to say that such a person is irresponsible, even evil. Cigarette argues we must finally conclude that he is in fact not a candidate for moral responsibility – that he is not a moral agent, not to be assigned prospective responsibilities, not to be held retrospectively responsible for his actions. In other words, it only makes sense to grade someone as responsible or irresponsible, so long as holding her responsible has any prospect of making her act more responsibly. The psychopath will never be responsive to blame, nor ever feel guilt. In fact, as someone who will never take any responsibility seriously, he does not qualify as a moral agent at all – as being responsible in its most basic sense. This might sound like writing the person a blank check to behave utterly immorally, but two points should be remembered: First, society protects itself against such people, often by incarcerating them as insane ("psychopathy" names a mental disorder). Second, the Kantian account reminds us that not to treat someone as responsible for

her actions is to fail to respect her as the author of her deeds. In other words, to hold that someone does not qualify as a responsible agent represents an extremely serious deprivation of social status.

Looking at the matter positively, we can also say that a person who exhibits the virtue of responsibility lives up to the three other aspects of responsibility in an exemplary way. First, she exercises the capacities of responsible moral agency to a model degree. Second, she approaches her previous actions and omissions with all due concern, being prepared to take responsibility for any failings she may have shown. And third, she takes her prospective responsibilities seriously, being both a capable judge of what she should do, and willing to act accordingly.