

4



Value and the Quest for the Good

There is beauty in sky and cloud and sea, in lilies and in sunsets, in the glow of bracken in autumn and in the enticing greenness of a leafy spring. Nature, indeed, is infinitely beautiful, and she seems to wear her beauty as she wears color or sound. Why then should her beauty belong to us rather than to her? Human character and human dispositions have value or worth, which belongs to them in the same sense as redness belongs to the cherry.

JOHN LAIRD, *A STUDY IN REALISM*

We never strive for, wish for, long for, or desire anything, because we deem it to be good, but, rather, we deem a thing good, because we strive for it, wish for it, long for it, or desire it.

BENEDICT DE SPINOZA, *ETHICS*

What sorts of things are valuable? Some items that we value are rather trivial, such as a new pair of shoes or one's preferred brand of soda. Yes we enjoy them, but they have no real urgency. Other things, though, seem to be of ultimate importance, and at the top of that list many of us would place the value of human life. After all, it's hard to find value in anything unless we're alive to experience it. Some of us might even claim to place an *absolute* value on human life. Now suppose I told you that I had invented a marvelous Convenience Machine that would save everyone an enormous amount of time and energy in our daily routines. However, the downside of the Convenience Machine is that its use would result in the deaths of over 75,000 Americans per year. Would you use this machine? Perhaps you'd refuse on the grounds that the value of life exceeds any amount of convenience.

But suppose our economy centered on the use of this machine, and without it, the nation would be thrown into an unparalleled economic depression. Perhaps you'd still refuse to use it and insist that we change our economic expectations rather than continually sacrifice so many lives.

Well, we in fact have this Convenience Machine in several brands: Chevrolet, Ford, Chrysler, Toyota, Honda, Mercedes, and so on. Motor vehicle accidents in the United States result in about 45,000 deaths a year; another 30,000 deaths are caused by diseases brought on by automobile pollution. So how much do we *really* value life? Perhaps not as much as we often claim, and we certainly do not value life as an absolute. Some people say that it is the quality of life rather than life itself that is valuable. The ancient Greeks and Romans believed that when life became burdensome, one had the obligation to commit suicide, for it was not the quantity of life that counted but the quality. As one Stoic philosopher put it, "Mere living is not a good, but to live well is a good."

The human life is just one example of a wide range of things that we find valuable, and a complete list of them would probably be impossible to create. Nicholas Rescher, though, classifies some basic values into these eight categories:¹

1. *Material and physical value*: health, comfort, physical security
2. *Economic value*: economic security, productiveness
3. *Moral value*: honesty, fairness, kindness
4. *Social value*: generosity, politeness, graciousness
5. *Political value*: freedom, justice
6. *Aesthetic value*: beauty, symmetry, grace
7. *Religious value*: piety, obedience, faith
8. *Intellectual value*: intelligence, clarity, knowledge

It's easy enough to devise a list of values like this: just think about what you do during the day and reflect on what is most important to you. What's less easy, though, is understanding why things are valuable to begin with and what, if anything, our various values have in common. In this chapter, we explore the notion of value and how value connects with issues of morality.

INTRINSIC AND INSTRUMENTAL VALUE

When we look at Rescher's list of basic values, we see that some seem to be valuable for their own sake, such as beauty and justice, while others are valuable because of their beneficial consequences, such as physical and economic security. The essential difference here is between intrinsic and instrumental goods. **Intrinsic goods** are good because of their nature and are not derived from other goods. By contrast,

instrumental goods are worthy of desire because they are effective means of attaining our intrinsic goods. Plato makes this distinction in his book, *The Republic*, where the characters Socrates and Glaucon are talking:

SOCRATES: Tell me, do you think there is a kind of good which we welcome not because we desire its consequences but for its own sake: joy, for example, and all the harmless pleasures which have no further consequences beyond the joy which one finds in them?

GLAUCON: Certainly, I think there is such a good.

SOCRATES: Further, there is the good which we welcome for its own sake and also for its consequences, knowledge, for example, and sight and health. Such things we somehow welcome on both accounts.

GLAUCON: Yes.

SOCRATES: Are you also aware of a third kind, such as physical training, being treated when ill, the practice of medicine, and other ways of making money? We should say that these are wearisome but beneficial to us; we should not want them for their own sake, but because of the rewards and other benefits which result from them.²

The question “What things are good or valuable?” is ambiguous. We need first to separate the kinds of values or goods there are. In the above, Socrates distinguishes three kinds of goods: (1) purely intrinsic goods (of which simple joys are an example); (2) purely instrumental goods (of which medicine and making money are examples); and (3) combination goods (such as knowledge, sight, and health), which are good in themselves *and* good as a means to further goods.

The essential difference is between intrinsic and instrumental goods. We consider some things good or worthy of desire (desirable) in themselves and other things good or desirable only because of their consequences. Intrinsic goods are good because of their nature. They are not derived from other goods, whereas instrumental goods are worthy of desire because they are effective means of attaining our intrinsic goods.

We may further distinguish an *instrumental good* from a *good instrument*. If something is an instrumental good, it is a means to attaining something that is intrinsically good; but merely to be a good instrument is to be an effective means to any goal, good or bad. For example, poison is a good instrument for murdering someone, but murder is not an intrinsically good thing; thus poison, in this use at least, is not an instrumental good.

Many things that we value are instrumental values. Socrates in our selection from *The Republic* mentions two instrumental values: medicine and money. Medicine is an instrumental good in that it can hardly be valued for its own sake. We can ask “What is medicine for?” The answer is, “It is to promote health.” But is health an intrinsic value or an instrumental one? Can we ask “What is health for?” Some will agree with Socrates that health is good for itself and for other things as well, such as happiness and creative activity. Others will dispute Socrates’ contention and judge health to be wholly an instrumental good.

Money is Socrates' other example of an instrumental value. Few, if any, of us really value money for its own sake, but almost all of us value it for what it can buy. When we ask "What is money for?" we arrive at such goods as food and clothing, shelter and automobiles, and entertainment and education. But are any of these really intrinsic goods, or are they all instrumental goods? When we ask, for example, "What is entertainment for?" what answer do we come up with? Most of us would mention enjoyment or pleasure, Socrates' example of an intrinsic good. Can we further ask "What is enjoyment or pleasure for?" We examine this question in the next section, but, before we do, we need to ask whether the notion of intrinsic values makes any sense.

Are there any intrinsic values? Are there any entities whose values are not derived from something else—that is, that are sought for their own sake, that are *inherently good*, good in themselves? Or are all values relative to desirers—that is, instrumental to goals that are the creation of choosers? Those who espouse the notion of intrinsic value usually argue that pleasure is an example of an intrinsic value and pain an example of an intrinsic disvalue: It is good to experience pleasure and bad to experience pain. Naturally, these philosophers admit that individual experiences of pleasure can be bad (because they result in some other disvalue such as a hangover after a drinking spree) and individual painful experiences can be valuable (for example, having a painful operation to save one's life). The intrinsicist affirms that pleasure is just better than pain. We can see this straight off. We do not need any arguments to convince us that pleasure is good or that gratuitous pain is intrinsically bad. Suppose we see a man torturing a child and order him to stop at once. If he replies, "I agree that the child is experiencing great pain, but why should I stop torturing her?" we would suspect some mental aberration on his part.

The nonintrinsicist denies that the preceding arguments have any force. The notion that the experience itself could have any value is unclear. It is only by our choosing pleasure over pain that the notion of value begins to have meaning. In a sense, all value is extrinsic, or a product of choosing. Many existentialists, most notably Jean-Paul Sartre, believe that we invent our values by arbitrary choice. The freedom to create our values and thus to define ourselves is godlike and, at the same time, deeply frightening, for we have no one to blame for our failures but ourselves. "We are condemned to freedom.... Value is nothing else but the meaning that you choose. One may choose anything so long as it is done from the ground of freedom."³

But this seems false. We do not choose most of our values in the same way we choose between two different majors or whether to have soup or salad with our meal. We cannot help valuing pleasure, health, happiness, and love and disvaluing pain and suffering. With regard to the fundamental values, they choose us, not we them. Even Sartre's condition for choosing a value, *freedom*, is not a value that we choose but have thrust upon us by our nature. We could override our freedom for other values, but we can no more choose whether to value it or not value it than we can choose whether or not to be hungry or thirsty after being deprived of food or drink for days. It is as though God or evolution pre-programmed us to desire these basic goods. And when we find someone who

does not value (or claims not to value) happiness, freedom, or love, we tend to explain this anomaly as a product of unfortunate circumstances.

THE VALUE OF PLEASURE

Philosophers divide into two broad camps: hedonists and nonhedonists. The hedonist (from *hedon*, Greek for “pleasure”) asserts that all pleasure is good, that pleasure is the only thing good in itself, and that all other goodness is derived from this value. An experience is good in itself if and only if it provides some pleasure. Sometimes, this definition is widened to include the lessening of pain, pain being seen as the only thing bad in itself. For simplicity’s sake, we will use the former definition, realizing that it may need to be supplemented by reference to pain.

Hedonists subdivide into two categories: (1) **sensualism**, the view that equates all pleasure with sensual enjoyment, and (2) **satisfactionism**, the view that equates all pleasure with satisfaction or enjoyment, which may not involve sensuality. Satisfaction is a pleasurable state of consciousness such as we might experience after accomplishing a successful venture or receiving a gift. The opposite of sensual enjoyment is physical pain; the opposite of satisfaction is displeasure or dissatisfaction.

The Greek philosopher Aristippus (ca. 435–366 BCE) espoused the sensualist position; that is, the only (or primary) good was sensual pleasure, and this goodness was defined in terms of its intensity.

This was also Mustapha Mond’s philosophy in Aldous Huxley’s *Brave New World*. The brave new world is a society of the future where people have been liberated from disease, violence, and crime through immunization, genetic engineering, and behavior modification. They are protected from depression and unhappiness through a drug, soma, that offers them euphoric sensations. Mustapha Mond, the brilliant manager of the society, defends this hedonistic utopia against one of the few remaining malcontents, the “Savage,” who complains that something of value is missing in this “utopia.” The following dialogue is between Mustapha Mond, the genius technocrat who governs the brave new world, and the malcontent, “Savage,” who believes that this hedonic paradise lacks something.

SAVAGE: Yes, that’s just like you. Getting rid of everything unpleasant instead of learning to put up with it. Whether ’tis better in the mind to suffer the slings and arrows of outrageous fortune, or to take arms against a sea of troubles and by opposing end them.... But you don’t do either. Neither suffer nor oppose. You just abolish the slings and arrows. It’s too easy.... Isn’t there something in living dangerously?

MUSTAPHA MOND: There’s a great deal in it.... Men and women must have their adrenals stimulated from time to time.... It’s one of the conditions of perfect health. That’s why we’ve made the VPS treatment compulsory.

SAVAGE: VPS?

MUSTAPHA MOND: Violent Passion Surrogate. Regularly once a month. We flood the whole system with adrenin. It's the complete physiological equivalent of fear and rage ... without any of the inconveniences.

SAVAGE: But I like the inconvenience.

MUSTAPHA MOND: In fact you're claiming the right to be unhappy.... Not to mention the right to grow old and ugly and impotent; the right to have syphilis and cancer; the right to have too little to eat; the right to live in constant apprehension of what may happen tomorrow; the right to be tortured by unspeakable pains of every kind.

SAVAGE (after
a long silence): I claim them all.

MUSTAPHA MOND
(shrugging
his shoulders): You're welcome.⁴

All but sensuously deprived adolescents (or those in a similar psychological state) would probably agree that the brave new world is lacking something. The sensuous version of pleasure is too simple.

Most hedonists since the third century BCE follow Epicurus (342–270 BCE), who had a broader view of pleasure:

It is not continuous drinkings and revellings, nor the satisfaction of lusts, nor the enjoyment of fish and other luxuries of the wealthy table, which produce a happy life, but sober reasoning, searching out the motives for all choice and avoidance, and banishing mere opinions, to which are due the greatest disturbance of the spirit.⁵

The distinction between pleasure as satisfaction and as sensation is important, and failure to recognize it results in confusion and paradox. One example of this is the paradox of masochism. How can it be that the masochist enjoys (that is, takes pleasure in) pain, which is the opposite of pleasure? "Well," the hedonist responds, "because of certain psychological aberrations, the masochist enjoys (as satisfaction) what is painful (as sensation)." But he or she does not enjoy (as sensation) what is painful (as sensation). There is also a two-level analysis to explain the masochist's behavior: On a lower, or basic, level, he is experiencing either pain or dissatisfaction, but on a higher level, he approves and finds satisfaction from that pain or dissatisfaction.

Nonhedonists divide into two camps: monists and pluralists. *Monists* believe that there is a single intrinsic value, but it is not pleasure. Perhaps it is a transcendent value, "the Good," which we do not fully comprehend but which is the basis of all our other values. This seems to be Plato's view. *Pluralists* generally admit that pleasure or enjoyment is an intrinsic good, but they add that there are other intrinsic goods as well, such as knowledge, friendship, aesthetic beauty, freedom, love, moral goodness, and life itself.

Hedonists such as Jeremy Bentham (1748–1832) argue that although these qualities are good, their goodness is *derived* from the fact that they bring pleasure or satisfaction. Such hedonists ask of each of the previously mentioned values, “What is it for?” What is knowledge for? If it gave no one any satisfaction or enjoyment, would it really be good? Why do we feel there is a significant difference between knowing how many stairs there are in New York City and whether or not there is life after death? We normally do not value knowledge of the first kind, but knowledge of the second kind is relevant for our enjoyment.

The hedonist asks, “What are friendship and love for?” If we were made differently and got no satisfaction out of love and friendship, would they still be valuable? Are they not highly valuable, significant instrumental goods because they bring enormous satisfaction?

Even moral commitment or conscientiousness is not good in itself, argues the hedonist. Morality is not intrinsically valuable but is meant to serve human need, which in turn has to do with bringing about satisfaction.

And, life certainly is not intrinsically good. It is quality that counts. An amoeba or a permanently comatose patient has life but no intrinsic value. Only when consciousness appears does the possibility for value arrive. Consciousness is a necessary but not a sufficient condition for satisfaction.

The nonhedonist responds that this is counterintuitive. Consider, for example, the possibility of living in a Pleasure Machine. We have invented a complex machine into which people may enter to find pure and constant pleasure. Attached to their brains will be electrodes that send currents to the limbic area of the cerebral cortex and other parts of the brain, producing very powerful sensations of pleasure. When people get into the machine, they experience these wonderful feelings. Would you enter such a machine?

If all you want is pleasure or satisfaction, then the Pleasure Machine seems the right choice. You’re guaranteed all the pleasure you’ve ever dreamed of—without frustration or competition from other people. But if you want to *do* something and *be* something (for example, have good character or a certain quality of personality) or experience reality (for example, friendship and competition), then you might think twice about this choice. Is the Pleasure Machine not just another addiction—like alcohol, heroin, cocaine, or crack? Once in the machine, would we become forever addicted to it? Furthermore, if all you want is pleasure, why not just hire someone to tickle you for a lifetime? Wouldn’t we become tired of being passive blobs—even if it was pleasurable? Most of us would reject such an existence as equivalent to that of a drugged cockroach.

Or suppose there were two worlds with the same number of people and the same amount of total pleasure, but in World I the people were selfish and even evil, whereas in World II the people were deeply moral. Wouldn’t it seem that World II was intrinsically better than World I?

Or imagine two lives, those of Suzy and Izzy. Suzy possesses 100 hedons (units of pleasure), even though she is severely retarded and physically disabled, whereas Izzy enjoys great mental acumen and physical prowess but has only 99 hedons. Isn’t it obvious that Izzy has the better life? But, hedonists are committed to saying that Suzy’s life is better, which seems implausible.

It was these sorts of cases that led John Stuart Mill (1806–1873, to be examined in Chapter 7)—in his classic work, *Utilitarianism*—to modify the hedonic doctrine, admitting that “it is better to be a human dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied.”⁶ He suggested that there were different qualities of pleasure and that those who had experienced the different kinds could distinguish among them. Whether the notion of *quality of pleasure* can save hedonism is a controversial matter, but many of us feel uneasy with the idea that pleasure alone is good. Some broader notion, such as *happiness* or *object of desire*, seems a more adequate candidate for what we mean by “value.”

ARE VALUES OBJECTIVE OR SUBJECTIVE ?

Do we desire the Good because it is good, or is the Good good because we desire it? The objectivist holds that values are worthy of desire whether or not anyone actually desires them; they are somehow independent of us. The subjectivist holds, to the contrary, that values are dependent on desirers, are relative to desirers.

The classic objectivist view on values (the absolutist version) was given by Plato (428–348 BCE), who taught that the Good was the highest form, ineffable, godlike, independent, and knowable only after a protracted education in philosophy. We desire the Good because it is good. Philosophers in the Platonic tradition hold to the independent existence of values apart from human or rational interest. For example, G. E. Moore claims that the Good is a simple, unanalyzable quality, such as the color yellow, but one that must be known through intuition. Moore believes that a world with beauty is more valuable than one that is a garbage dump, regardless of whether there are conscious beings in those worlds:

Let us imagine one world exceedingly beautiful. Imagine it as beautiful as you can ... and then imagine the ugliest world you can possibly conceive. Imagine it simply one heap of filth.⁷

Moore asks us whether, even if there were no conscious being who might derive pleasure or pain in either world, we would prefer the first world to exist rather than the second. Moore believes that it is obvious that the beautiful world is inherently better, but the objector asks, “What good is such a world if there is no one (even God) to enjoy it?”

Other, weaker objectivist versions treat values as emergent properties, or qualities in the nature of things. That is, just as the wetness of water is not in the H₂O molecules but in the interaction of our nervous system with millions of those molecules, and just as smoothness is not in the table that I am touching but in the relationship between the electrical charges of the subatomic particles of which the table is made up and my nervous system, so values (or good qualities) emerge in the relationship between conscious beings and physical and social existence. They are synergistic entities, depending on both our nature and their objective properties.

For example, if we were not beings with desires, we would not be in a position to appreciate values; but once there are such beings, certain things—such as pleasure, knowledge, freedom, friendship, and health—will be valuable, and others—such as pain, suffering, boredom, loneliness, disease, and death—will be disvalued or not valued for their own sake. This synergistic view recognizes both a subjective and an objective aspect to value.

Subjectivism treats values as merely products of conscious desire. The American pragmatist Ralph Barton Perry (1876–1957) states that a value is simply the object of interest.⁸ Values are created by desires, and they are valuable just to that degree to which they are desired: The stronger the desire, the greater the value. The difference between the subjectivist and the weak objectivist position (or mixed view) is simply that the subjectivist makes no normative claims about “proper desiring,” instead judging all desires as equal. Anything one happens to desire is, by definition, a value, a good.

The objectivist responds that we can separate the Good from what one desires. We can say, for example, that Joan desires more than anything else to get into the Pleasure Machine, but it is not good; or that John desires more than anything else to join the Satanic Society, where he will pursue evil for evil’s sake, engaging in sado-masochistic behavior, but it is not good (not even for John). There is something just plain bad about the Pleasure Machine and the Satanic Society, even if Joan and John never experience any dissatisfaction on account of them.

On the other hand, suppose Joan does not want to have any friends and John does not want to know any history, literature, philosophy, or science (beyond whatever is necessary for his needs as a devotee of hardcore pornography or mud wrestling). The objectivist would reply that it really would be an objectively good thing if Joan did have friends and if John knew something about history, literature, philosophy, and science.

Perhaps a way to adjudicate the disagreement between the subjectivist and the objectivist is to imagine an Ideal Desirer, a person who is impartial and has maximal knowledge of the consequences of all actions. What the Ideal Desirer chooses is by definition the “good,” and what he or she disdains is the “bad.” If so, we can approximate such an ideal perspective by increasing our understanding and ability to judge impartially. The study of philosophy, especially moral philosophy, has as one of its main goals such an ability.

THE RELATION OF VALUE TO MORALITY

Typically, value theory is at the heart of moral theory. The question, however, is whether moral right and wrong are themselves intrinsic values (as Kant states, the moral law is “a jewel that shines in its own light”) or whether rightness and wrongness are defined by their ability to further nonmoral values such as pleasure, happiness, health, and political harmony. To begin to understand this question and to get an overview of the workings of morality, let me offer a schema of the moral process (Figure 4.1), which may help in locating the role of values in moral theory.

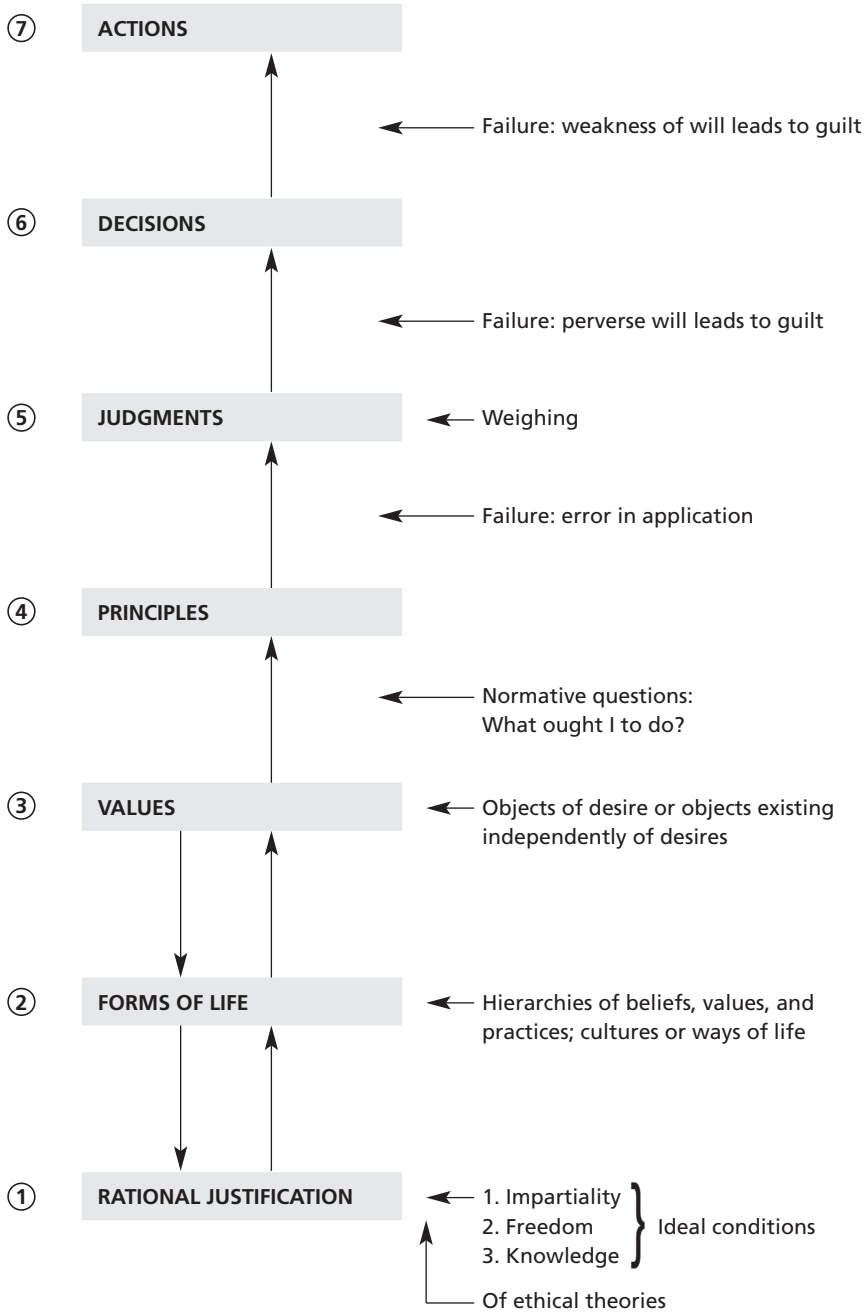


FIGURE 4.1 Schema of the moral process

The location of values in the schema of the moral process (box 3) indicates that values are central to the domain of morality. They are the source of principles (box 4) and rooted in the forms of life (box 2). Examples of values are life, loving relationships, freedom, privacy, happiness, creative activity, knowledge, health, integrity, and rationality. From our values, we derive principles (box 4), which we may call action-guiding value “instantiators” or “exemplifiers” (because they make clear the action-guiding or prescriptive force latent in values). From the value “life,” we derive the principles “Promote and protect life” and/or “Thou shall not kill.” From the value “freedom,” we derive the principle “Thou shall not deprive another of his or her freedom.” From the value “privacy,” we derive the principle “Respect every person’s privacy.” From the value “happiness,” we derive the principle “Promote human happiness,” and so forth with all the other values.

This schema makes no judgment as to whether values are objective or subjective, intrinsic or instrumental. Neither does it take a stand on whether values or principles are absolute; they need not be absolute. Most systems allow that all or most values and principles are overridable. That is, they are considerations that direct our actions, and whenever they clash, an adjudication must take place to decide which principle overrides the other in the present circumstances.

We often find ourselves in moral situations in which one or more principles apply. We speak of making a judgment as to which principle applies to our situation or which principle wins out in the competition when two or more principles apply (box 5). The correct principle defines our duty. For example, we have the opportunity to cheat on a test and immediately judge that the principle of honesty (derived from the value integrity) applies to our situation. Or there might be an interpersonal disagreement in which two or more people differ on which of two values outweighs the other in importance, as when Mary argues that Jill should not have an abortion because the value of life outweighs Jill’s freedom and bodily integrity, but John argues that Jill’s freedom and bodily integrity outweigh the value of life.

Even after we judge which principle applies, we are not yet finished with the moral process. We must still *decide* to do the morally right act. Then finally, we must actually *do* the right act.

Note the possibilities for failure all along the way. We may fail to apply the right principle to the situation (the arrow between boxes 4 and 5). For example, we may simply neglect to bring to mind the principle against cheating. This is a failure of application. But even after we make the correct judgment, we may fail to make the right choice, deciding to cheat anyway. In this case, we have a perverse will (the arrow between boxes 5 and 6). Finally, we may make the correct choice but fail to carry out our decision (the arrow between boxes 6 and 7). We call this *weakness of will*: We mean to do the right act but simply are too morally weak to accomplish it. In our example, we meant to refrain from cheating but couldn’t control ourselves. “The good that I would, I do not, but the evil that I would not, that I do.”⁹

A more controversial matter concerns the deep structure in which values are rooted. Some theories deny that there is any deep structure but assert instead that values simply exist in their own right—independently, as it were. More often, however, values are seen as rooted in whole forms of life (box 2) that can be

actual or ideal, such as Plato's hierarchical society or Aristotle's aristocracy or the Judeo-Christian notion of the kingdom of God (the ideal synagogue or church). Ways of life or cultures are holistic and hierarchical combinations of beliefs, values, and practices.

The deepest question about morality is whether and how these forms of life are justified (box 1). Are some forms of life better or more justified than others? If so, how does one justify a form of life? Candidates for justification are ideas such as God's will, human happiness, the flourishing of all creation, the canons of impartiality and knowledge, a deeply rational social contract (Hobbes and Rawls), and the like. For example, a theist might argue that the ideal system of morality (that is, the ideal form of life) is justified by being commanded by God. A utilitarian would maintain that the ultimate criterion is the promotion of welfare or utility. A naturalist or secular humanist might argue that the ideal system is justified by the fact that it best meets human need or promotes human flourishing or that it would be the one chosen by ideally rational persons. Some ethicists would make level 2 the final source of justification, denying that there is any ideal justification at all. These are the ethical relativists, who contend that each moral system is correct simply by being chosen by the culture or individual.

The main point of the schema, however, is not to decide on the exact deep structure of morality but to indicate that values are rooted in cultural constructs and are the foundation for moral principles upon which moral reasoning is based. We could also devise a similar schema for the relationship between values and virtues (to be discussed in Chapter 9). Each virtue is based on a value and each vice on a disvalue.

THE GOOD LIFE

Finally, we want to ask what kind of life is most worth living. Aristotle (384–322 BCE) wrote long ago that what all people seek is happiness:

There is very general agreement; for both the common person and people of superior refinement say that it is happiness, and identify living well and doing well with being happy; but with regard to what happiness is they differ, and the many do not give the same account as the wise. For the former think it is some plain and obvious thing, like pleasure, wealth or honor.¹⁰

What is happiness? Again, the field divides up among objectivists, subjectivists, and combination theorists. The objectivists, following Plato and Aristotle, distinguish happiness from pleasure and speak of a single ideal for human nature; if we do not reach that ideal, then we have failed. Happiness (from the Greek *eudaimonia*, literally meaning “good demon”) is not merely a subjective state of pleasure or contentment but the kind of life we would all want to live if we understood our essential nature. Just as knives and forks and wheels have functions, so do species, including the human species. Our function (sometimes

called our “essence”) is to live according to reason and thereby to become a certain sort of highly rational, disciplined being. When we fulfill the ideal of living the virtuous life, we are truly happy.

Plato speaks of happiness as “harmony of the soul.” Just as the body is healthy when it is in harmony with itself and the political state is a good state when it is functioning harmoniously, so the soul is happy when all its features are functioning in harmonious accord, with the rational faculty ruling over the spirited and emotional elements. Although we no doubt know when we are happy and feel good about ourselves, the subjective feeling does not itself define happiness, for people who fail to attain human excellence can also feel happy via self-deception or ignorance.

The objectivist view fell out of favor with the rise of the evolutionary account of human nature, which undermined the sense of a preordained essence or function. Science cannot discover any innate *telos*, or goal, to which all people must strive. The contemporary bias is in favor of value pluralism—that is, the view that there are many ways of finding happiness: “Let a thousand flowers bloom.” This leads to subjectivism.

The subjectivist version of happiness states that happiness is in the eyes of the beholder. You are just as happy as you think you are—no more, no less. The concept is not a descriptive one but a first-person evaluation. I am the only one who decides or knows whether I am happy. If I feel happy, I am happy, even though everyone else despises my lifestyle. Logically, happiness has nothing to do with virtue, although—due to our social nature—it usually turns out that we will feel better about ourselves if we are virtuous.

The combination view tries to incorporate aspects of both the objectivist and the subjectivist views. One version is John Rawls’s “plan of life” conception of happiness: There is a plurality of life plans open to each person, and what is important is that the plan be an integrated whole, freely chosen by the person, and that the person be successful in realizing his or her goals. This view is predominantly subjective in that it recognizes the person as the autonomous chooser of goals and a plan. Even if a person should choose a life plan

whose only pleasure is to count blades of grass in various geometrically shaped areas such as park squares and well-trimmed lawns, ... our definition of the good forces us to admit that the good for this man is indeed counting blades of grass.¹¹

However, Rawls recognizes an objective element in an otherwise subjective schema. There are primary goods that are necessary to any worthwhile life plan: “rights and liberties, powers and opportunities, income and wealth, ... self-respect, ... health and vigor, intelligence and imagination.”¹² The primary goods function as the core (or the hub of the wheel) from which may be derived any number of possible life plans (the spokes). But unless these primary goods (or most of them) are present, the life plan is not an authentic manifestation of an individual’s autonomous choice of his or her own selfhood. Thus, it is perfectly possible that people believe themselves to be happy when they really are not.

Although subjectivist and plan-of-life views dominate the literature today, there is some movement back to an essentialist, or Aristotelian, view of happiness as a life directed toward worthwhile goals. Some lifestyles are more worthy than others, and some may be worthless. Philosopher Richard Kraut asks us to imagine a man who has as his idea of happiness the state of affairs of being loved, admired, or respected by his friends and who would hate to have his “friends” only pretend to care for him. Suppose his “friends” really do hate him but “orchestrate an elaborate deception, giving him every reason to believe that they love and admire him, though in fact they don’t. And he is taken in by the illusion.”¹³ Can we really call this man happy?

Or suppose a woman centers her entire life around an imaginary Prince Charming. She refuses to date—let alone marry—perfectly eligible young men; she turns down educational travel opportunities lest they distract her from this wonderful future event; for ninety-five years, she bores all her patient friends with tales of the prince’s imminent appearance. As death approaches at age ninety-six, after a lifetime of disappointment, she discovers that she’s been duped; she suddenly realizes that what appeared to be a happy life was a stupid, self-deceived, miserable existence. Would we say that our heroine was happy up until her deathbed revelation? Do these thought experiments not indicate that our happiness depends, at least to some extent, on reality and not simply on our own evaluation?

Or suppose we improve on our Pleasure Machine, turning it into a Happiness Machine. This machine is a large tub that is filled with a chemical solution. Electrodes are attached to many more parts of your brain. You work with the technician to program all the “happy experiences” that you have ever wanted. Suppose that includes wanting to be a football star, a halfback who breaks tackles like a dog shakes off fleas and who has a fondness for scoring last-minute game-winning touchdowns. Or perhaps you’ve always wanted to be a movie star and to bask in the public’s love and admiration. Or maybe you’ve wanted to be the world’s richest person, living in the splendor of a magnificent castle, with servants faithfully at your beck and call. In fact, with the Happiness Machine you can have all of these plus passionate romance and the love of the most beautiful (or handsome) people in the world. All these marvelous adventures would be simulated, and you would truly believe you were experiencing them. Would you enter the Happiness Machine?

What if I told you that once you were unplugged, you could either stay out or go in for another round but that no one who entered the machine ever chose to leave of his or her own accord, having become addicted to its pleasures and believing that reality could never match its ecstasy. Now you have an opportunity to enter the Happiness Machine for the first time. Will you enter? If not, are you not voting against making the subjectivist view (or even the plan-of-life view) the sole interpretation of happiness?

When I ask this question in class, I get mixed responses. Many students say they would enter the Happiness Machine; most say they would not. I myself would not, for the same reason that I do not use drugs and rarely watch television or spectator sports—because some very important things are missing that are necessary for the happy life. What are these vital missing ingredients?

1. *Action.* We are entirely passive in the machine, a mere spectator. But the good life requires participation in our own destiny. We don't just want things to happen to us; we want to accomplish things, even at the risk of failure.
2. *Freedom.* Not only do we want to do things, but we want to make choices. In the Happiness Machine, we are entirely determined by a preordained plan—we cannot do otherwise. In fact, we cannot do anything but react to what has been programmed into the machine.
3. *Character.* Not only do we want to do things and act freely, but we also want to *be* something and someone. To have character is to be a certain kind of person, ideally one who is trustworthy, worthy of respect, and responsible for one's actions. In the machine, we lose our identity. We are defined only by our experience but have no character. We are not persons who act out of set dispositions, for we never act at all. We are mere floating blobs in a glorified bathtub.
4. *Relationships.* There are no real people in our Happiness Machine life. We subsist in splendid solipsism. All the world is a figment of our imagination as dictated by the machine; our friends and loved ones are mere products of our fancy. But we want to love and be loved by real people, not by phantasms.

In sum, the Happiness Machine is a myth, all *appearance* and no *reality*—a bliss bought at too high a price, a deception! If this is so and if reality is a necessary condition for the truly worthwhile life, then we cannot be happy in the Happiness Machine. But neither can we be happy outside of the Happiness Machine when the same necessary ingredients are missing: activity, freedom, moral character, loving relationships, and a strong sense of reality.

The objective and subjective views of happiness assess life from different perspectives, with the objectivist assuming that there is some kind of independent standard of assessment and the subjectivist denying it. Even though there seems to be an immense variety of lifestyles that could be considered intrinsically worthwhile or happy and even though some subjective approval or satisfaction seems necessary before we are willing to attribute the adjective “happy” to a life, there do seem to be limiting conditions on what may count as happy. We have a notion of *fittingness* for the good life, which would normally *exclude* being severely retarded, being a slave, or being a drug addict (no matter how satisfied) and which would *include* being a deeply fulfilled, autonomous, healthy person. It is better to be Socrates dissatisfied than to be the pig satisfied, but only the satisfied Socrates is happy.

This moderate objectivism is set forth by John Stuart Mill. Happiness, according to Mill, is

not a life of rapture; but moments of such, in an existence made up of few and transitory pains, many and various pleasures, with a decided predominance of the active over the passive, and having as the foundation of the whole, not to expect more from life than it is capable of bestowing.¹⁴

This conception of happiness is worth pondering. It includes activity, freedom, and reality components, which exclude being satisfied by the passive experience in the Happiness Machine, and it supposes (the context tells us this) that some pleasing experiences are better than others. I would add to Mill's definition the ingredients of moral character and loving relations. A closer approximation might go like this:

Happiness is a life in which there exists free action (including meaningful work), loving relations, and moral character and in which the individual is not plagued by guilt and anxiety but is blessed with peace and satisfaction.

The *satisfaction* should not be confused with complacency; rather, it means contentment with one's lot—even as one strives to improve it. Whether this neo-objectivist, Millian view of happiness is adequate, you must decide.

CONCLUSION

In this chapter, we've seen that there is a range of ways to dissect the notion of moral goodness. Some goods are intrinsic because of their nature and are not derived from other goods, and others are instrumental because they are effective means of attaining intrinsic goods. Goods are often connected with pleasure; sensualism equates all pleasure with sensual enjoyment whereas satisfactionism identifies all pleasure with satisfaction or enjoyment, which may not involve sensuality. There is a debate whether values are objective or subjective. Plato held the former position, maintaining that goods have an independent existence of values apart from human or rational interest; Perry held the latter view that values are merely products of conscious desire. Although value theory is at the center of moral theory, there is dispute about whether the moral notions of right and wrong are themselves intrinsic values. Finally, there is the issue of how values are connected with human happiness and the good life, particularly whether there is a human purpose, or *telos*, that defines our capacity for happiness in terms of specific values.

NOTES

1. Nicholas Rescher, *Introduction to Value Theory* (Prentice Hall, 1969), p. 16.
2. Plato's *The Republic*, Bk. II, trans. G. M. A. Grube (Hackett, 1980).
3. Jean-Paul Sartre, *Existentialism and Human Emotions*, trans. Bernard Frechtman (Philosophical Library, 1957), pp. 23, 48–49.
4. Adapted from Aldous Huxley, *Brave New World* (Harper & Row, 1932), pp. 286–287.

CHAPTER TWO

Meet the Values: Intrinsic, Final & Co.

2.1 Introduction

To the extent that we justify actions and attitudes in terms of some value they produce or express, we do not normally need to reach for fundamental, basic, or ultimate values. If a certain medicine promises to be effective against my headache, and accordingly I take it (or you advise me to take it), we need not question whether a headache-free condition would be good in itself, or because of some other good state related to it – say, having the capacity to think clearly. If a certain policy is thought to reduce unemployment or improve working conditions, we can support it without asking why having a decent job is a good thing: maybe work, in certain conditions, is good for its own sake, or maybe it is good because of other good things to which it connects – financial autonomy, or the possibility of achievement.

But certainly there are plenty of contexts where our attention focuses on a deeper dimension of value. For instance, many of us would take a critical attitude towards a businessman intent on accumulating money exclusively for the sake of it, quite regardless of how money impacts on other aspects of his life. The feeling is that money, however important it may be, is the sort of thing to be valued not for its own sake, but for the sake of something else; at any rate, it cannot be the *only* thing to be valued for its own

sake. Conversely, someone who appears to treat their own children exclusively as potential workforce for the family would be the target of blame: whether or not people can have some instrumental value, few of us believe that one's own children can be regarded exclusively as potential workforce. In these and similar examples, our criticism is based on an intuitive distinction between what is valuable for its own sake (one's own children) and what is instrumentally valuable (money).

Similar distinctions are traditionally made by philosophers in their search for the basic elements of a good life, and for ultimate standards of action: the attempt is to pick out what, in some sense, matters fundamentally, as opposed to what matters for the sake of something else. The aim of this chapter is to make distinctions between value concepts which are often not distinguished, either within philosophy or in ordinary value talk: final value, exclusively final value, unconditional value, intrinsic value, necessary/essential value. The idea is to alert the reader to the importance of keeping these concepts separate. On the argumentative plane, I provide some defence for the idea that something can be valuable for its own sake (i.e. finally valuable, the sense in which one's children are valuable and money arguably is not), and yet doesn't *have to* be unconditionally valuable, nor *must* it owe its value to its intrinsic properties. This idea is worth exploring, since it has been assumed by most philosophers that what is valuable for its own sake can only be valuable in itself, that is, in isolation from other things, and not conditionally on other values. Final value, according to this tradition, must be *the* stopping place of our evaluations, or is not really 'final'. I will argue that this tradition, which has a distinguished history and has recently been revived, considerably restricts the ways in which it is appropriate to value things, and depends on a questionable assumption on what it means to be 'final'. In section 2.6 I also explain why we should distinguish merely instrumental extrinsic value from other forms of extrinsic value.

2.2 Final and unconditional value: Some philosophical examples

When philosophers build their substantive axiologies, i.e. their theories of what is good and bad, they generally seek to identify what is ultimately good and bad. In stating their views, they use various phrases. One is ‘good for its own sake’. For instance, after an examination of people’s opinions and values, Aristotle thus describes the sense in which *eudaimonia* (happiness as a state of flourishing) is the ultimate good:

If, then, there is some end of the things we do, which we desire for its own sake (everything else being desired for the sake of this), and if we do not choose everything for the sake of something else (for at that rate the process would go on to infinity, so that our desire would be empty and vain), clearly this must be the good and the chief good ... Now such a thing happiness, above all else, is held to be; for this we choose always for self and never for the sake of something else, but honour, pleasure, reason, and every virtue we choose indeed for themselves (for if nothing resulted from them we should still choose each of them), but we choose them also for the sake of happiness, judging that by means of them we shall be happy. Happiness, on the other hand, no one chooses for the sake of these, nor, in general, for anything other than itself. (*Nicomachean Ethics* I 2, 7)¹

Another commonly found phrase is ‘good in itself’, equally used by otherwise very different philosophers such as Kant and Bentham. Kant applies it to the good will, i.e. willing or intending to act according to the moral law:

It is impossible to conceive anything at all in the world, or even out of it, which can be taken as good without qualification, except a *good will* ... A good will is not good because of what it effects or accomplishes – because of its fitness for attaining some

¹Aristotle is obviously not the first philosopher to draw such distinctions. For example, see Plato, *Republic* 357.

proposed end; it is good through its willing alone – that is, good in itself. (Kant 1964: 61–2)

Bentham proposes instead a hedonist view of value:

Strictly speaking, nothing can be said to be good or bad, but either in itself, which is the case only with pain or pleasure; or on account of its effects, which is the case only with things that are the causes or preventives of pain and pleasure. (Bentham 1970: 11)

Yet another phrase is ‘good/desirable as an end’, for instance to be found in Mill’s statement of hedonism (here misleadingly identified with utilitarianism, which includes but does not reduce to a theory of the good):

The utilitarian doctrine is, that happiness is desirable, and the only thing desirable, as an end; all other things being only desirable as means to that end. (Mill 1861: IV 2)

How are we sure that Aristotle, Kant, Bentham, and Mill are talking about the same thing? First, despite different terminologies, *eudaimonia*, pleasure, good will, or happiness are all contrasted with things that are ‘good for the sake of something else’, that is to say, things that are good because they stand in a certain (typically but not necessarily) instrumental relation to the former items. For example, from a hedonist point of view, health is good insofar as it produces, facilitates, or simply contains experiences of pleasure, or prevents painful experiences. Devoid of such relationships to pain and pleasure, health would be neither good nor bad. Second, they all mean to pick out what can be described as *final values*. Since in value theory this term is not always understood in the same way, here is the idea we are working with. Something is finally valuable when it is valuable for its own sake, that is, as suggested in Chapter 1, when we have reason or it is fitting to favour or value it for its own sake. The idea is fully normative: something that is finally valuable deserves, as such, a certain kind of treatment.

However, beyond this shared concern for final values in this sense, other notions are mixed in the quotations above, and it is instructive to tease them apart. One is the concept of something

that is not only finally valuable, but *exclusively finally valuable*: all the value it has or could ever have is final value. This is what after all for Aristotle sets *eudaimonia* apart from things like reason and virtue, which *are* finally valuable but not *exclusively* finally valuable, since they are valuable also for contributing to other things, such as happiness. Also Kant is explicit about this, as he claims that the usefulness of the good will could not add any value to it (1964: 62). And the same goes for hedonists: of course a particular experience of pleasure could give us further future pleasures, and to that extent be instrumentally as well as finally good, but there is no non-pleasure-related state which might make an experience of pleasure additionally valuable in this way. Pleasure could not be worth seeking for reasons other than pleasure: it could only be good for its own sake. So, alongside contrasting final to non-final values, these philosophers meant to pick out exclusively vs. non-exclusively final values.

It is important to note, however, that the search for exclusively final value is not inevitable. Those who hold a *pluralist* axiology, whereby at least two things are finally valuable – e.g. pleasure and virtue – would have no problem admitting that virtue is both finally valuable, and also valuable when and because it produces pleasure, and vice versa for pleasure. One such example is W. D. Ross's axiology, which includes innocent pleasure, knowledge, and virtue as final values (2002 [1930]). Exclusively final value thus is a notion that only *monist* axiologies require: if there is no other final value than V, then there is nothing else for the sake of which V could be valuable. Of course, if you are a pluralist, you *might* still believe that each of your chosen two (or more) final values is exclusively finally good: e.g. that the value of virtue cannot be enhanced by pleasure or vice versa. But the monist *needs to* believe in a value that is exclusively final.

A second notion implicit in the quotations above is that those final values are also *unconditional* values, in the sense that their value does not presuppose the value of anything else. Again, if you are monist, e.g. you think that pleasure is the only final value around, then it follows that the value of pleasure cannot depend on the value of anything else. For an example of a *conditional* value, Kant famously mentions happiness: someone's happiness is valuable only if they are worthy of it (due to their good will). In this sense, Kant denies that happiness *as such* is valuable for

its own sake. It is a problematic question whether Kant thereby counts as a monist, since in some interpretations it appears that now *two* things are *finally* valuable: good will (unconditionally), and deserved happiness (conditionally on the value of good will).

The possibility of such an interpretation already suggests that final value *need not* be unconditional. For a different example, consider a valuable relationship such as friendship. Friends are usually disposed to help each other out, and they take pleasure or derive happiness from their relationship. Plausibly, if their happiness and altruistic dispositions had no value, their friendship would have no value either. In this sense, the value of friendship is not unconditional. But this does not mean that we are to value friendship simply for the sake of these other things (happiness, altruistic action). Similarly, the value of altruistic action itself can plausibly be said to presuppose the goodness of happiness and the badness of suffering: if suffering were not bad, there would be no merit in alleviating others' suffering. But this, again, does not mean that altruism is to be admired and encouraged simply for the sake of minimizing suffering. In sum, goods like friendship, altruism, or happiness (in Kant's view) can be finally valuable and yet (unlike Aristotle's *eudaimonia* or Kant's good will) not unconditionally so.

To prevent possible misunderstandings: of course, even values that are in this sense unconditional (i.e. not conditional on other values) may depend on the obtaining of certain *enabling conditions*; e.g. nobody can have a good will if they are not sufficiently mentally sophisticated creatures. Being a mentally sophisticated creature is a necessary condition for someone to have a good will, and therefore for the value of the good will to be instantiated. Moreover, plausibly nobody can be a mentally sophisticated creature in isolation from others (say, their parents): there being other people around is another necessary condition for the value of good will. But this is a dependence on *factual*, rather than *evaluative*, conditions: enabling conditions of this sort do not refer to the *value* of being mentally sophisticated or the value of there being people around. In this sense, an unconditional final value can and normally does depend on factual conditions.

If we might be tempted to see the presence of *evaluative* enabling conditions as undermining the status of a given value as final (wrongly, as I have briefly suggested), no similar temptation should

arise with regard to *factual* enabling conditions. Consider another example. Any experience we may have depends on us being alive: dead people have no experiences. Therefore also the value of our experiences (e.g. our pleasures) depends on us being alive. But it does not follow that experiences are valuable for the sake of being alive. Rather, being alive appears valuable, if at all, for the sake of the experiences it makes possible. The value of being alive (as an enabling condition) is therefore non-final (see Chapter 5 for further discussion of enabling conditions).²

2.3 Intrinsic value and final value

The phrases ‘for its own sake’ and ‘in itself’ naturally suggest a further consideration: what makes these things finally valuable must somehow be found ‘in’ the valuable thing, rather than in something totally or partially ‘outside’ it. This is why such values have historically been called *intrinsic*, and therefore contrasted with *extrinsic* ones. G. E. Moore so defines intrinsic value: ‘To say that a kind of value is ‘intrinsic’ means merely that the question whether a thing possesses it, and in what degree it possesses it, depends solely on the intrinsic nature of the thing in question’ (1993 [1922]: 286). And in order to determine what has intrinsic value, he devised the so-called isolation test: ‘it is necessary to consider what things are such that, if they existed *by themselves*, in absolute isolation, we should yet judge their existence to be good’ (1993 [1903]: 236). Now, certain things just cannot exist in absolute isolation: e.g. a state of pleasure necessarily has a subject and (at least for some sorts of pleasures) an object (what it is pleasure *at*). So the test seems unserviceable as it stands: we cannot coherently imagine a world containing states of pleasures but no subject and no object of those pleasures. However, in the light of Moore’s definition of intrinsic value, we can propose an amended isolation test: in order to determine the intrinsic value of *x*, we have

²Henry Sidgwick makes it clear that the factual or at least conceptual dependence of beauty and knowledge on minds is compatible with favouring beauty and knowledge as final ends – though he eventually does not take them to be so (1981: 114).

to focus on x 's intrinsic nature alone, and *on that basis* see if we judge x to be good – presumably, this is something we can do with pleasures and other states.³

For instance, Kant's good will is said to keep its value regardless of any relation to actions it may or may not originate. For a classical hedonist, it is the intrinsic qualities of pleasures (such as their sensory quality, intensity, or duration) which determine their value: if I feel good because I think I have won the lottery, then my state has the same intrinsic value whether in fact I have or have not won the lottery. And intrinsic value can belong, of course, also to goods, e.g. the beauty of a work of art, whose 'intrinsic nature' is a matter of certain relations among different parts (say, colours and shapes being in a certain arrangement): such relations make up the work of art for what it is. In this sense, they are 'internal' as opposed to external relations, such as the fact that the work of art provokes emotional experiences to viewers, or that it is a unique piece in its genre. Such external relations do not affect the intrinsic value of the object as a beautiful work of art. As for Moore himself, the isolation test delivered the result that the enjoyment of personal relationships and the appreciation of beauty are, not the only intrinsic values, but the highest ones.

But is all *final* value intrinsic value? So did Moore assume. And until not long ago, such a question might have sounded otiose. If something is to be valued for *its own* sake, then it seemed obvious that what makes it worth valuing in such a way must be found in the intrinsic nature of the object. And if what makes something worth valuing is to be found partly 'outside' the object itself, then it seems that the value of the object cannot be final. However, precisely these statements should already make it clear that final/non-final and intrinsic/extrinsic value are, at least, conceptually distinguishable pairs: final/non-final refers to the correct or appropriate way of valuing something, while intrinsic/extrinsic refers to the metaphysical location of the good-making properties of something. So a Moorean view which equates final and intrinsic value need not be obviously true.

In particular, many believe that final values can be extrinsic:

³See Lemos's 'intentional isolationism' (1994: 10–11).

objects which are valuable for *their own* sake partly thanks to their relations to *other* objects. This can happen in many ways.⁴

- 1 One is where final value depends on the value of other things. As suggested above, if friends' happiness were not valuable, and friendship did not contribute to friends' happiness – both facts being conceived of as externally related to the friendship – probably friendship would not be valuable for its own sake. For a more concrete example: a particular fur coat might be regarded as valuable for its own sake, as an outstanding piece of handicraft, yet so only assuming an appropriate evaluative background. If fur coats were not in general instrumentally valuable for protection against the cold they provide, this particular coat could not have any value, let alone any final value.⁵
- 2 Another case is where final value might be enhanced by an object's relational properties: a fine work of art, beautiful and thus already valuable on account of its intrinsic features, might have its value increased by its being a unique or rare piece (a kind of relational property, since it implies the absence of *other* things, or many other things, like it). And such uniqueness or rarity need not make the

⁴See Korsgaard (1983), Kagan (1998), Rabinowicz and Rønnow-Rasmussen (1999), Langton (2007). Korsgaard, who basically single-handedly started this line of thought, went on to argue for an ultimate dependence of final value on the extrinsic property of being valued by a rational agent. But this claim of hers is best understood as a general meta-ethical thesis about the nature of value, rather than as the denial that there is any real intrinsic value (see section 1.6). Being valued by a rational agent, after all, is not to be counted among the good-making properties on top of whatever intrinsic or extrinsic properties already make an object (e.g. a fur coat) finally valuable. In other words, intrinsic value need not be metaphysically mind-independent value (this is a confusion which arguably Moore himself caused in his 1922 essay).

⁵Jonathan Dancy would claim that if the enabling conditions (factual or evaluative) for something to be a value are extrinsic properties, but the ground of value (the good-making feature) is an intrinsic property, then the value is still intrinsic (2004a: 172). In the example, friendship could still be an intrinsic value, even if happiness is an extrinsic enabling condition for its value. While I accept the distinction between enabling conditions and grounds, and hence the possibility that intrinsic (and final) value is variable (see Chapter 5), I prefer for illustrative purposes to use 'intrinsic value' in the way defined by Moore, whose notion of 'dependence' admittedly doesn't discriminate between enabling conditions and grounds of value.

object additionally valuable simply in instrumental terms, say because it increases its economic value, or because it provides particular pleasure to viewers or owners.

- 3 Yet another case is where the final value of an object does not just counterfactually depend on the value of something else, but results from or is *exhaustively* grounded on some relation to other objects. Uniqueness or rarity can give an otherwise unremarkable item (say, a stamp) a new final value. A certain type of car might have the ability to race at unusually high speed: this relational, indeed causal, feature may conceivably make the car valuable (e.g. worth maintaining) for its own sake. Were it not for this feature, the car would simply have the instrumental, non-final value that most other cars have. Or consider Napoleon's hat: a dull, worn-out accessory, but it is not absurd to believe it worth preserving for its own sake simply because it belonged to such an extraordinary person. In each of these cases, of course, there is also a long story to be told about counterfactual evaluative dependences. For instance, in the case of the car, we are dealing with something like the value of pushing physical boundaries; in the case of the hat, the value of keeping traces of historically crucial figures, and in turn the value of history. So in these cases, final value is extrinsic in two ways: first, it is grounded in relational properties of the object, and second, such relational properties make the object finally valuable only on the condition of something else than the object itself being valuable.

If these examples are persuasive, then final value can be extrinsic. It also follows that we cannot always apply Moore's isolation test to discover what has final value. Remember that we had to imagine the object in isolation from any other thing. Then we would miss the final value of Napoleon's hat, since we would ignore who the hat belonged to, and likewise the final value of a rare stamp, since we would ignore how many other stamps like it are around. We would miss the additional value of a unique artwork, since we wouldn't know that there are no other exemplars of that. We would miss the final value of the fur coat, since we would have to ignore the point of producing and wearing fur coats. And we would possibly mistake the final value of any given friendship,

since we wouldn't know if it is actually good for the friends – we would ignore an external condition for its final value.

2.4 The reduction to facts

Given the plethora of examples available, it seems that the burden of proof lies on those who take final value to be always intrinsic value, i.e. value which depends only on an object's intrinsic features. How could one hope to show that? One strategy starts from the question: what are the real bearers of value, i.e. what sorts of entities really have value? Most of the examples assume that *individual objects* (coats, hats, stamps ...) are the sort of entities that can have final value. The strategy invites us to reconsider this assumption. It seems that we could, in principle, always understand or translate the putative extrinsic final value of an individual object in terms of the final value of a *fact*, or a *state of affairs*, which includes the object together with the relevant relational properties. For example, the final value of Napoleon's hat can be seen as really the final value of the fact that

(F1) *there is a hat which belonged to Napoleon.*

Alternatively, it could be seen as the final value of the fact that

(F2) *this hat belonged to Napoleon.*

Either F1 or F2 would be both worth valuing for its own sake, and in virtue of their intrinsic or internal properties: that the hat belonged to Napoleon is indeed what these facts are all about, be it F1 or F2. So this translation strategy, or move from objects to facts, would give the result that a Moorean needs: a picture where final value is always intrinsic. Similar moves can be made, in principle, with other putative cases of extrinsic final value.⁶

⁶See Zimmerman (2001: chapter 3), who explicitly advocates this move. But the general idea goes back to Ross (2002 [1930]: 112), and has been endorsed by many value theorists since (Chisholm (2005 [1972]), Lemos (1994: 20–31). See Rabinowicz and Rønnow-Rasmussen (1999, 2003) for discussion and criticism.

It is interesting to note that the F2-style translation, which avoids the existential commitment, is perhaps preferable to F1 for two reasons. First, it is not obvious that to value something for its own sake always means to value its *existence*. This might work for artefacts which are worth preserving (and thus keeping into existence), but what about wishing a painless death to our sick dog as a way of valuing him for his own sake? Second, precisely this point shows that in general we value the existence of x only if we value x in the first place. F1-style translation instead sees valuing x 's existence as a prior condition on valuing x , so it seems to get things the wrong way around.

But why adopt such a reduction or translation manoeuvre in the first place? Why can't we be happy with final values that are extrinsic? One immediate reason might be simplicity. Extrinsic final values give us a less unified picture of how value comes into existence. And, as we have seen, we wouldn't be able to use Moore's isolation test, which might have given us a handy procedure to look for the things to care about for their own sake.

A more theoretical reason, suggested by Michael Zimmerman (2005: 194), is that if there are long stories to be told about, e.g. why Napoleon's hat is a valuable thing, i.e. stories that must refer to values other than the value of the hat itself, then what we have is at most a *derivative* value. But final values, in principle, should rather play the role of 'endpoint values': once we reach them, no helpful explanation of them can be given, except saying that they are good 'as such' – i.e. 'for *their* own sake'. That's why extrinsic values of the sort indicated above cannot be *final* values for Zimmerman: explanation (or rather justification) does not stop at them. Only values that are intrinsic can play such an endpoint role, because there's no looking beyond them to explain why they are valuable. And it seems that only *facts*, and not individual objects, can be guaranteed to play this double role of carrying final, i.e. non-derivative, and intrinsic value.

Moreover, the reduction to facts does correspond to a natural way of talking about value, so it's not an ad hoc or merely technical move. We obviously do not only evaluate objects or persons, but also facts or situations: we say that it is good *that* this and that happened, or that it would be better *if* that did not happen, and so on. And it also matches with the idea that many of our fitting responses to value are what are generally called *propositional attitudes*, such as desire or preference, e.g. desiring *that the war end*. If, as claimed in Chapter 1,

values must be the sort of things towards which fitting responses can be directed, then it is easy to see why values at least *could* consist in proposition-like entities, such as facts or states of affairs.

However, the reduction manoeuvre is not easily accomplished. I will mention two challenges. First, since the strategy has to be carried out for the whole spectrum of final values, its proponents are committed to the claim that *only* propositional attitudes are fitting responses to final value. If final values are all facts such as F2, then it seems that attitudes, such as respect, preservation, or certain forms of love, which are thing- or person-oriented, rather than fact-oriented, cannot be appropriate responses to final values. We respect and love individual people (or animals), and this is a different attitude than loving *that such people have this or that characteristic*. We want to preserve Napoleon's hat, rather than the fact that it belonged to Napoleon – indeed, preserving the latter seems to make almost no sense. According to the reduction manoeuvre, thing- or person-oriented attitudes might still be in some sense fitting to their objects, but they won't match with final value: final value will rather belong to the sorts of states which we can value by a relevant propositional attitude. So the fittingness of preserving Napoleon's hat will reduce, it seems, to the fittingness of valuing or cherishing F2, the fact that the hat belonged to Napoleon. While this latter example may give some plausibility to the strategy (after all preserving the hat *is* a way of valuing F2), it is not equally clear that, e.g. respect and love for individual persons can *always* be seen as fitting because of the fittingness of a kind of respect and love towards some fact about these persons. In a sentence, the reduction of finally valuable things to finally valuable facts commits one to a reduction of fitting responses (to final value) to fact-oriented attitudes, and the latter reduction might be hard to sustain, given the variety in the kind of responses that we normally identify as 'valuing for its own sake'. Perhaps then bearers of final value are just as various as our responses: 'We value many different kinds of things, including at least the following: objects and their properties (such as beauty), persons, skills and talents, states of character, actions, accomplishments, activities and pursuits, relationships, and ideals' (Scanlon 1998: 95).⁷

⁷See Anderson (1993) for a view at the other extreme: states of affairs are *never* fundamental bearers of final value. Only individuals (in particular, people) are.

Second, the theoretical justification suggested by Zimmerman for regarding all final value as intrinsic is weak. There is no initial reason to think that final values, understood as things worth valuing for their own sake, also need to play the role of non-derivative or endpoint values, in the sense that no explanation of their value can be given, except pointing to their own intrinsic nature. Conceptually, we are again dealing with different notions: final value, which refers to the appropriate way of responding to something, and non-derivative value, a structural-metaphysical notion. Consider again friendship. The value of friendship, and thus the respect a particular friendship deserves from third parties, can be partly explained by reference to other values: the mutual desire for happiness that friends typically have is good; the happiness thus achieved is also good; and, at least on some views, friends must make each other happy by and large in morally permissible ways. The value of friendship then is not an 'endpoint' value, because appreciating the value of any given friendship is not independent of appreciating the value of other things (happiness, morality). But valuing, e.g. respecting, a particular friendship on the condition that it doesn't involve immoral conduct, obviously does not mean valuing it for the sake of something else, namely moral permissibility. One can recognize a value as conditional on other values, and in this sense 'derivative', and still fittingly care about it for its own sake.

What does seem to be true is that, if a final value is derivative or conditional on other values, then this reflects on the appropriateness of responses. If I care about Napoleon's hat for its own sake, but show otherwise utter indifference towards Napoleon's sword, or indeed to objects belonging to comparable historical figures (say, Sitting Bull's war bonnet), then my valuing is not of the fitting kind, precisely because it is insensitive to the structure of the final value towards which it is directed (assuming, of course, that the hat is only valuable for its historical connection). But such interrelation among fitting attitudes does not subtract anything from the final value of the object – it is not as if its value gets 'thinned out' in the network of attitudes required towards similar objects.

In the attempt to reduce final values to intrinsic, non-derivative values we can see again the concern with *unconditional* values which occupied Aristotle and Kant. Moreover, if such attempt is carried out by the reduction manoeuvre of construing the value

of individual entities as the value of facts including those entities, what we get are *essential* values, i.e. values which depend on the essential properties of the valuable object. It is an essential property of F2 that it concerns Napoleon's hat and its having belonged to Napoleon: a different hat or a different owner would make up a different fact. In general, a fact like 'that x is P ' has the essential property of concerning x , or ascribing P to x . If a fact is valuable because of such intrinsic property, then it is also essentially valuable. In this sense, final values turn out to be incorruptible, that is, they remain constant for as long as their bearer remains what it is. It is a good question, to say the least, whether it is a defensible consequence of such a view that only unconditional and essential values are really worth valuing for their own sake, whereas values that are had conditionally or contingently do not deserve that same kind of response.⁸

2.5 Intrinsic and conditional value

Having defended the possibility of final extrinsic values, we can now also clarify the relations between intrinsic/extrinsic and unconditional/conditional value. Let's consider first unconditional value. Unconditionally good things do not depend on the value of anything else. Therefore whatever it is that makes them good must be looked for among their intrinsic properties: in their intrinsic nature. As said, even unconditional values might depend on *factual* enabling conditions: the unconditional value of pleasure depends on the subject being alive. Now, such enabling conditions can be intrinsic or extrinsic properties of the valuable object; e.g. there being other people around might be a factual *extrinsic* condition for the value of a good will. But this doesn't mean that therefore the good will has an unconditional *extrinsic* value. This is because Moore's definition of intrinsic value and the amended isolation test for it *presuppose* that whatever necessary conditions for x to exist

⁸Bradley 2006 suggests that 'Mooreans' (for whom final value is intrinsic and necessary) and their opponents (whom he calls 'Kantians') might be talking past each other, and simply have in mind two different concepts of value. I discuss this in Chapter 7.

and be intelligible as such, be they intrinsic or extrinsic properties of x , have already been counted in. So, for instance, it would make no sense to apply Moore's test to the good will, and conclude that the good will has no *intrinsic* value because one instance of the good will in a world without other agents would strike us as valueless. Rather, when we apply the test, we should imagine a world where necessary conditions for a good will to exist and be intelligible as such are taken for granted (such as the existence of other people besides the agent). Factual dependence on external conditions doesn't make a value extrinsic.

What about conditional value? Here a certain object x is dependent for its value on the value 'of something else'. Now, this 'something else' can be intrinsically or extrinsically related to x . Earlier I considered the value of a friendship, and suggested that it may depend not only on whether it produces happiness for both friends, but on whether such happiness is indeed good for them or more generally valuable. Happiness, and its value, are in this way *extrinsically* related to any given friendship: when we apply Moore's test on friendship, its producing valuable happiness or not is one of those facts we should abstract from, because it is not a necessary condition for any friendship to exist – friendship can produce misery as well as happiness. So the value of friendship appears to be conditional *and* extrinsic (but not for these reasons non-final, as argued above).

On the other hand, if the 'something else' is intrinsically related to x , then x has conditional but intrinsic value. A good example is offered by Thomas Hurka's theory of virtues, in which virtues are attitudes of this form: 'loving (desiring, etc.) the good for its own sake' and 'hating (avoiding, etc.) the bad for its own sake' (Hurka 2001). Virtues are therefore second-order final values. Their structure makes their value at once *conditional*, because it presupposes the first-order value of the objects of love and hate, but also *intrinsic*, since it is an intrinsic feature of a virtue that it involves an attitude towards the good or the bad.

An instructive contrast between intrinsic and extrinsic conditional value emerges from two different readings of Kant's claims about deserved happiness. If we take the valuable state of affairs to be 'S is happy and deserves being happy (i.e. has a good will)', then its value is conditional on the value of the good will, and yet intrinsic, since S's having a good will is a part of the state of affairs:

it is intrinsically related to it. On the other hand, if the valuable state of affairs is simply ‘S is happy’, then its value is conditional on S’s having a good will, but extrinsic, because S’s having a good will is externally related to *this* state of affairs.

The crucial point from this discussion is that intrinsic value can be conditional value. This is significant, since it is easy to confuse the two notions: it is natural to wonder, how can something’s value be intrinsic, belong to ‘the thing itself’, if it depends on the value of something else? But just like we shouldn’t assume final value to be necessarily intrinsic, so should we not assume that intrinsic value is always unconditional.⁹

2.6 Elimination of extrinsic value?

There is a lingering concern with extrinsic values. If x owes its value to the value of something else y , then the worry is that x really has no value at all. Some philosophers are tempted to draw this conclusion when considering extrinsic value of the instrumental sort. Thus Ross on the value of acts: ‘Whatever value [an act] has independently of its motive is instrumental value, i.e. *not goodness at all*, but the property of producing something that is good’ (2002 [1930]: 133, my emphasis). There is a sense in which the act does not contribute any value to the world. In computing how much value the world contains, we are not going to add instrumental value *on top of* whatever value the act has caused. A world where the same valuable states of affairs occur through other means would contain the same amount of value (other things being equal). Hence the act is good-causing, but not literally good.

⁹To complete the picture: exclusively final values can be conditional values. For instance, Hurka’s virtues are attitudes whose final value is conditional on the value of their objects. But Hurka could argue that virtues are *only* valuable for their own sake. Can exclusively final values be extrinsic? Consider some of our examples. Friendship’s value depends on the value of happiness, but this by itself doesn’t mean that necessarily it is fitting to favour friendship (also) for the sake of happiness. Rather it means that one cannot value friendship without also valuing happiness. Of course, plausibly friendship can be favoured for its beneficial consequences, so probably it is not an exclusively final value. Presumably deserved happiness, for Kant, is a value that is both extrinsic (given the second reading in the main text) and exclusively final (it could only be worth favouring for its own sake).

Similarly, we may call a medicine healthy, but the truth is that it is health-restoring (there is nothing healthy about a pill-shaped chemical aggregate). What is healthy is the state you are in upon taking the medicine.

But could we draw *eliminativist* conclusions for all extrinsic value? For instance, some philosophers speak of 'signatory value': an X-ray is signatorily good if it indicates something else that is good (e.g. that the tumour has gone). (X-rays of course are in general also instrumentally good, as aids to medical knowledge.) Or there can be 'contributory value': a certain motif in a painting is good in this sense if it contributes to the (aesthetic) value of the painting as a whole. It is good 'as a part'. It would be tedious to recount all forms of extrinsic value, because it would require drawing a list of all relevant relations: causing, being a sign of, being a part of a whole, being historically connected to (as in Napoleon's hat), etc. But the question now is: if we grant that one sort of relation (instrumental) to value means that an object (an act, say) *really* is valueless, why not generalize and conclude that the concept of extrinsic value is, at best, a handy way of talking about value, but does not capture a genuine evaluative reality? And if extrinsic value doesn't capture a genuine evaluative reality, then a fortiori all final value must be intrinsic.

While not everyone agrees with Ross, it might be worth ending this chapter by suggesting why, even if he is right about instrumental value, not all extrinsic value can be eliminated like that. The argument for elimination assumes that what has instrumental value does not contribute any value over and above the value of its causal consequences. However, this does not seem to apply to the cases of extrinsic final values considered above.

First, it is true that Napoleon's hat has no value apart from its historical connection. But Napoleon's hat contributes value precisely in being an exemplar of a supposedly valuable category of things – objects (maybe of a certain kind) that belonged to important historical figures. Destroy the hat, and you have directly reduced the amount of value contributed by this category. On the other hand, once you imagine away acts and in general things that only have instrumental value – while keeping constant the amount of valuable states of affairs they would otherwise produce – you will not have diminished in the slightest the amount of value in that world. And this is after all what makes sense of why it is

appropriate to value the former for *its own* sake, and the latter for the sake of something else.

Second, regarding values such as friendship which are conditional on external factors, the elimination argument would work if it were true that the fact of friendship would contribute *no* value, over and above the value contributed by other facts which are conditions for friendship's value (e.g. mutual happiness produced within a friendship). But it is hard to see how to show this much. While it might be held that a given friendship loses in positive value if it makes one or both friends worse off, *what* loses in value is the complex of expectations, mutual feelings, shared history, etc., which define friendship for what it is, and which – prior to philosophical arguments to the contrary – constitute its final value and its specific valuable contribution 'to the amount of value in the world'. Of course some philosophers (e.g. hedonists) would be ready to make the *substantive* claim that friendship is only instrumentally valuable, e.g. insofar as it promotes the general happiness. But this move would not be acceptable, since here we were looking at possible reasons why extrinsic values are eliminable (and a fortiori non-final) *qua* extrinsic. It seems that no general argument for elimination can be found.

2.7 Summary

In this chapter I have distinguished several pairs of value concepts:

- final vs. non-final value: what is fitting to favour for its own sake vs. what is fitting to favour for the sake of something else;
- exclusively final vs. non-exclusively final value: what is fitting to favour *only* for its own sake vs. what is fitting to favour for its own sake and for the sake of something else;
- unconditional vs. conditional value: what is fitting to favour (for its own sake) independently of whether it is fitting to favour something else vs. what is fitting to favour (for its own sake or not) not independently of whether it is fitting to favour something else;

- intrinsic vs. extrinsic value: what is fitting to favour because of its intrinsic properties vs. what is fitting to favour (for its own sake or not) partly because of its extrinsic properties;
- essential/necessary vs. contingent value: what is fitting to favour in all possible worlds where it occurs vs. what is fitting to favour in some but not all possible worlds where it occurs.

I have argued that we should keep an open mind as to how these concepts relate to one another. In particular, I have suggested that final value need not be unconditional value, and criticized a strategy to equate intrinsic and final value based on the idea that only facts or states of affairs are bearers of final value. Since the resulting view is that there can be final extrinsic values, then I had to defuse a worry that the very category of extrinsic value could be dispensed with.