I. INTRODUCTION: DEFINITION AND ROOTS

Nonconsequentialism is a type of normative ethical theory that denies that the rightness or wrongness of our conduct is determined solely by the goodness or badness of the consequences of our acts or of the rules to which those acts conform. Nonconsequentialism does not deny that consequences can be a factor in determining the rightness of an act. It does insist that even when the consequences of two acts or act-types are the same, one might be wrong and the other right. Hence, nonconsequentialism denies the truth of both act and rule consequentialism, which are understood as holding that the right act or system of rules is the one that maximizes the balance of good consequences over bad ones as determined by an impartial calculation of goods and bads.\(^1\) (Henceforth, I shall refer to this as “maximizing the good.”) This sort of consequentialist calculation requires that we have a theory of what is good and bad; it may be an extremely liberal theory, holding that killings are bad or that autonomy is good, but we are still required to maximize the good.\(^2\) ...

Contemporary nonconsequentialism finds its spiritual roots in the work of Immanuel Kant and W. D. Ross. Some nonconsequentialists are especially drawn to Kant’s second formulation of the Categorical Imperative, which specifies that we should always treat rational humanity in oneself and in others as an end-in-itself and never merely as a means, and to his distinction between perfect and imperfect duties. Persons are said to have a special kind of unconditional value—value independent of serving anyone’s (even their own) ends and independent of their being in a particular context—that makes them worthy of respect. Merely counting each person’s interests in a consequentialist calculation of overall good, while it seems to literally distinguish persons from mechanical tools, is not enough to ensure that we treat someone as an end-in-itself in the Kantian sense. Rather, it is thought, if I am an end-in-myself then this fact can constrain even conduct that would maximize overall good. ...

\(^1\) Consequentialism can also be joined with a nonimpartial conception of the good; for example, the right act is the one that maximizes my good. I shall not deal with such forms of consequentialism in the text.

\(^2\) Some consequentialists allow for “satisficing” the good. That is, producing enough but not maximizing. I shall not consider this view in the text. Satisficing theories can also conflict with nonconsequentialist principles, since the latter can deny that the rightness of conduct is solely a function of producing enough of a balance of good consequences over bad.
Despite the importance of Kantian theory to nonconsequentialists, some question whether an act is impermissible just because the agent treats someone as a mere means or not as an end-in-himself in doing the act. For suppose I do an act that is justified by its great good consequences despite some foreseen side effect harm to a few people. However, I do the act not for its good consequences, but only in order to produce the harm to the people as an end in itself (and I would not in any other way constrain myself for their sakes). Presumably, this still does not make my doing the act impermissible.

The second element of Kant’s legacy that appeals to some contemporary nonconsequentialists is his distinction between perfect and imperfect duties. The perfect duties describe specifically what we must do and they take precedence over the imperfect duties, which give us leeway in how or when we fulfill them. Thus, in Kant’s view, I may not kill one person in order to fulfill a duty to save others. Contemporary nonconsequentialists, however, often diverge from Kant’s absolutist conception of perfect duties (i.e., that such duties always take precedence over imperfect duties), and some claim merely that the class of negative duties (e.g., not to harm) is more stringent than the class of positive duties (e.g., to aid). Some contemporary interpreters of Kant have argued that his theory is not absolutist and does not imply, for example, that lying to someone in order to stop him from committing a murder is wrong.\(^3\) (It is not clear, however, that this interpretation, which seems to rely on the view that people’s bad acts can lead to their forfeiting a right not to be lied to, can also yield the truth that it is permissible to lie to *an innocent bystander*, if this is necessary to stop a murderer.)

W. D. Ross is another major inspiration for contemporary nonconsequentialism. Although Ross thought that there was a prima facie duty of beneficence, he also thought that there are numerous other prima facie duties, for example, a duty not to harm, a duty of gratitude, and a duty to do justice. If these prima facie duties conflict, as he thought they might, we have no single scale on which to weigh them or rule by which to order them so as to determine what our actual duty is. In this sense, the duties are incommensurable, but that need not mean that conflicts between them cannot be correctly decided by the exercise of judgment.\(^4\) Some contemporary nonconsequentialists have tried to modify Ross’s view by more precisely determining the relative weights or ordering of prima facie duties, or by more precisely characterizing the prima facie duties, so that it becomes clearer which takes precedence. This latter endeavor might require stating duties so that they specify their own limits or finding more basic duties than the ones Ross described that do not as easily come into conflict with each other.

---


II. CONTEMPORARY NONCONSEQUENTIALISM OUTLINED

Nonconsequentialism is now typically thought to include prerogatives [often called ‘options’] not to maximize the good and constraints on producing the good. A prerogative [i.e., option] denies that agents must always maximize good consequences. Hence, it allows for the possibility that some acts are supererogatory, these being acts that, though they are not morally required, are morally valuable, sometimes in virtue of producing better consequences. Constraints limit what we may do in pursuit of our own, or even the impartial, good. The most commonly proposed constraints are a strong duty not to harm (contrasted with a weaker duty to aid) and/or a strong duty not to intend harm (contrasted with a weaker duty not to cause or allow harm that is merely foreseen). Those who are only partially nonconsequentialists might advocate prerogatives but no constraints or constraints but no prerogatives.

However, commonly proposed constraints ignore important moral complexities. Consider, for example, the constraint on harming. In the Trolley Case, a runaway trolley will kill five people, if a bystander does not divert it onto another track where, he foresees, it will kill one person. Nonconsequentialists typically think that the bystander may divert the trolley—killing one person to save the five—although, in other cases, they oppose killing one person to save five. An appropriate constraint might better capture nonconsequentialist judgments of cases. If it does, it will capture the precise way in which an individual is thought to be inviolable and protected by a negative right not to be harmed, even if the harm would help to maximize the good. (Saying that someone is inviolable is a bit stronger than just saying that he has a right not to be harmed, as some rights might be permissibly infringed and then a person with such an infringeable right would not be inviolable to the same degree.)

Many nonconsequentialists employ a distinctive methodology. They test and develop theories or principles by means of intuitive judgments about cases. They compare the implications that proposed principles of permissible conduct have for hypothetical cases (such as the Trolley Case) with their considered judgments about what can permissibly be done in such cases. If the implications of the principles and judgments conflict, they may develop alternative principles. If the implications of the principles and judgments are compatible, the nonconsequentialist must still offer a theory identifying the fundamental, morally significant factors that underlie the principles in order for those principles to be fully justified. ...

---

5 For example, Samuel Scheffler, *The Rejection of Consequentialism* (New York: Oxford University Press, 1982).
... Finally, there is one imperative which directly commands certain conduct without making its condition some purpose to be reached by it. This imperative is categorical. It concerns not the material of the action and its intended result, but the form and principle from which it originates. What is essentially good in it consists in the mental disposition, the result being what it may. This imperative may be called the imperative of morality ... (belonging to free conduct as such, i.e., to morals). ...

There is ... only one categorical imperative. It is: Act only according to that maxim by which you can at the same time will that it should become a universal law. ... [T]he universal imperative of duty can be expressed as follows: Act as though the maxim of your action were by your will to become a universal law of nature.

We shall now enumerate some duties, adopting the usual division of them into duties to ourselves and to others and into perfect and imperfect duties.

A man who is reduced to despair by a series of evils feels a weariness with life but is still in possession of his reason sufficiently to ask whether it would not be contrary to his duty to himself to take his own life. Now he asks whether the maxim of his action could become a universal law of nature. His maxim, however is: For love of myself, I make it my principle to shorten my life when by a longer duration it threatens more evil than satisfaction. But it is questionable whether this principle of self-love could become a universal law of nature. One immediately sees a contradiction in a system of nature whose law would be to destroy life by the feeling whose special office is to impel the improvement of life. In this case it would not exist as nature; hence that maxim cannot obtain as a law of nature, and thus it wholly contradicts the supreme principle of all duty.

Another man finds himself forced by need to borrow money. He well knows that he will not be able to repay it, but he also sees that nothing will be lent him if he does not firmly promise to repay it at a certain time. He desires to make such a promise, but he has enough conscience to ask himself whether it is not improper and opposed to duty to relieve his distress in such a way. Now, assuming he does decide to do so, the maxim of his action would be as follows: When I believe myself to be in need of money, I will borrow money and promise to repay it, although I know I shall never be able to do so. Now this principle of self-love or of his own benefit may very well be compatible with his whole future welfare, but the question
is whether it is right. He changes the pretension of self-love into a universal law and then puts the question: How would it be if my maxim became a universal law? He immediately sees that it could never hold as a universal law of nature and be consistent with itself; rather it must necessarily contradict itself. For the universality of a law which says that anyone who believes himself to be in need could promise what he pleased with the intention of not fulfilling it would make the promise itself and the end to be accomplished by it impossible; no one would believe what was promised to him but would only laugh at any such assertion as vain pretense. …

[Another] man, for whom things are going well, sees that others (whom he could help) have to struggle with great hardships, and he asks, “What concern of mine is it? Let each one be as happy as heaven wills, or as he can make himself; I will not take anything from him or even envy him; but to his welfare or to his assistance in time of need I have no desire to contribute.” If such a way of thinking were a universal law of nature, certainly the human race could exist, and without doubt even better than in a state where everyone talks of sympathy and good will or even exerts himself occasionally to practice them while, on the other hand, he cheats when he can and betrays or otherwise violates the right of man. Now although it is possible that a universal law of nature according to that maxim could exist, it is nevertheless impossible to will that such a principle should hold everywhere as a law of nature. For a will which resolved this would conflict with itself, since instances can often arise in which he would need the love and sympathy of others, and in which he would have robbed himself, by such a law of nature springing from his own will, of all hope of the aid he desires.

The foregoing are a few of the many actual duties, or at least of duties we hold to be actual, whose derivation from the one stated principle is clear. We must be able to will that a maxim of our action become a universal law; this is the canon of the moral estimation of our action generally. Some actions are of such a nature that their maxim cannot even be thought as a universal law of nature without contradiction, far from it being possible that one could will that it should be such. In others this internal impossibility is not found, though it is still impossible to will that that maxim should be raised to the universality of a law of nature, because such a will would contradict itself. We easily see that a maxim of the first kind conflicts with stricter or narrower (imprescriptable) duty, that of the latter with broader (meritorious) duty. Thus all duties, so far as the kind of obligation (not the object of their action) is concerned, have been completely exhibited by these examples in their dependence upon the same principle. …

But suppose that there were something the existence of which in itself had absolute worth, something which, as an end in itself, could be a ground of definite laws. In it and only in it could lie the ground of a possible categorical imperative (i.e., of a practical law).
Now, I say, man and, in general, every rational being exists as an end in himself and not merely as a means to be arbitrarily used by this or that will. In all his actions, whether they are directed toward himself or toward other rational beings, he must always be regarded at the same time as an end. All objects of inclination have only conditional worth, for if the inclinations and needs founded on them did not exist, their object would be worthless. The inclinations themselves as the source of needs, however, are so lacking in absolute worth that the universal wish of every rational being must be indeed to free himself completely from them. Therefore, the worth of any objects to be obtained by our actions is at times conditional. Beings whose existence does not depend on our will but on nature, if they are not rational beings, have only relative worth as means, and are therefore called "things"; rational beings, on the other hand, are designated “persons” because their nature indicates that they are ends in themselves (i.e., things which may not be used merely as means). Such a being is thus an object of respect, and as such restricts all [arbitrary] choice. Such beings are not merely subjective ends whose existence as a result of our action has a worth for us, but are objective ends (i.e., beings whose existence is an end in itself). Such an end is one in the place of which no other end, to which these beings should serve merely as means, can be put. Without them, nothing of absolute worth could be found, and if all worth is conditional and thus contingent, no supreme practice principle for reason could be found anywhere.

Thus if there is to be a supreme practical principle and a categorical imperative for the human will, it must be one that forms an objective principle of the will from the conception of that which is necessarily an end for everyone because it is an end in itself. Hence this objective principle can serve as a universal law. The ground of this principle is: rational nature exists as an end in itself. Man necessarily thinks of his own existence in this way, and thus far it is a subjective principle of human actions. Also every other rational being thinks of his existence on the same rational ground which holds also for myself; thus it is at the same time an objective principle from which, as a supreme practical ground, it must be possible to derive all laws of the will. The practical imperative, therefore, is the following: Act so that you treat humanity, whether in your own person or in that of another, always as an end and never as a means only. ...
What Makes Right Actions Right?
from The Right and the Good
by W.D. Ross (1930)

The point at issue is that to which we now pass, viz. whether there is any general character which makes right acts right, and if so, what it is. ...

The view that what produces the maximum pleasure is right has for its bases the views (1) that what produces the maximum good is right, and (2) that pleasure is the only thing good in itself. If they were not assuming that what produces the maximum good is right, the utilitarians’ attempt to show that pleasure is the only thing good in itself, which is in fact the point they take most pains to establish, would have been quite irrelevant to their attempt to prove that only what produces the maximum pleasure is right. If, therefore, it can be shown that productivity of the maximum good is not what makes all right actions right, we shall a fortiori have refuted hedonistic utilitarianism.

When a plain man fulfils a promise because he thinks he ought to do so, it seems clear that he does so with no thought of its total consequences, still less with any opinion that these are likely to be the best possible. He thinks in fact much more of the past than of the future. What makes him think it right to act in a certain way is the fact that he has promised to do so—that and, usually, nothing more. That his act will produce the best possible consequences is not his reason for calling it right. What lends colour to the theory we are examining, then, is not the actions (which form probably a great majority of our actions) in which some such reflection as ‘I have promised’ is the only reason we give ourselves for thinking a certain action right, but the exceptional cases in which the consequences of fulfilling a promise (for instance) would be so disastrous to others that we judge it right not to do so. It must of course be admitted that such cases exist. If I have promised to meet a friend at a particular time for some trivial purpose, I should certainly think myself justified in breaking my engagement if by doing so I could prevent a serious accident or bring relief to the victims of one. And the supporters of the view we are examining hold that my thinking so is due to my thinking that I shall bring more good thereby than by not doing so. A different account may, however, be given of the matter, an account which will, I believe, show itself to be the true one. It may be said that besides the duty of fulfilling promises I have and recognize a duty of relieving distress, and that when I think it right to do the latter at the cost of not doing the former, it is not because I think I shall produce more good thereby but because I think it the duty which is in the circumstances more of a duty. This account surely corresponds much more closely with what we really think in such a situation. If, so far as I can see, I could bring equal amounts of good

---

8 These are not strictly speaking duties, but things that tend to be our duty, or prima facie duties.
into being by fulfilling my promise and by helping someone to whom I had made no promise, I should not hesitate to regard the former as my duty. Yet on the view that what is right is right because it is productive of the most good I should not so regard it.

There are two theories, each in its way simple, that offer a solution of such cases of conscience. One is the view of Kant, that there are certain duties of perfect obligation, such as those of fulfilling promises, of paying debts, of telling the truth, which admit of no exception whatever in favor of duties of imperfect obligation, such as that of relieving distress. The other is the view of, for instance, Professor Moore and Dr. Rashdall, that there is only the duty of producing good, and that all ‘conflicts of duties’ should be resolved by asking ‘by which action will most good be produced?’ But it is more important that our theory fit the facts than that it be simple, and the account we have given above corresponds (it seems to me) better than either of the simpler theories with what we really think, viz. that normally promise-keeping, for example, should come before benevolence, but that when and only when the good to be produced by the benevolent act is very great and the promise comparatively trivial, the act of benevolence becomes our duty.

In fact the theory of ‘ideal utilitarianism’, if I may for brevity refer so to the theory of Professor Moore, seems to simplify unduly our relations to our fellows. It says, in effect, that the only morally significant relation in which my neighbors stand to me is that of being possible beneficiaries by my action. They do stand in this relation to me, and this relation is morally significant. But they may also stand to me in the relation of promisee to promiser, of creditor to debtor, of wife to husband, of child to parent, of friend to friend, of fellow countryman to fellow countryman, and the like; and each of these relations is the foundation of a prima facie duty, which is more or less incumbent on me according to the circumstances of the case. When I am in a situation, as perhaps I always am, in which more than one of these prima facie duties is incumbent on me, what I have to do is to study the situation as fully as I can until I form the considered opinion (it is never more) that in the circumstances one of them is more incumbent than any other; then I am bound to think that to do this prima facie duty is my duty sans phrase in the situation.

I suggest ‘prima facie duty’ or ‘conditional duty’ as a brief way of referring to the characteristic (quite distinct from that of being a duty proper) which an act has, in virtue of being of a certain kind (e.g. the keeping of a promise), of being an act which would be a duty proper if it were not at the same time of another kind which is morally significant. Whether an act is a duty proper or actual duty depends

---

9 Some will think it, apart from other considerations, a sufficient refutation of this view to point out that I also stand in that relation to myself, so that for this view the distinction of oneself from others is morally insignificant.
on all the morally significant kinds it is an instance of. The phrase ‘prima facie duty’ must be apologized for, since (1) it suggests that what we are speaking of is a certain kind of duty, whereas it is in fact not a duty, but something related in a special way to duty. Strictly speaking, we want not a phrase in which duty is qualified by an adjective, but a separate noun. (2) ‘Prima’ facie suggests that one is speaking only of an appearance which a moral situation presents at first sight, and which may turn out to be illusory; whereas what I am speaking of is an objective fact involved in the nature of the situation, or more strictly in an element of its nature, though not, as duty proper does, arising from its whole nature. ...

There is nothing arbitrary about these prima facie duties. Each rests on a definite circumstance which cannot seriously be held to be without moral significance. Of prima facie duties I suggest, without claiming completeness or finality for it, the following division.¹⁰

(1) Some duties rest on previous acts of my own. These duties seem to include two kinds, (a) those resting on a promise or what may fairly be called an implicit promise, such as the implicit undertaking not to tell lies which seems to be implied in the act of entering into conversation (at any rate by civilized men), or of writing books that purport to be history and not fiction. These may be called the duties of fidelity. (b) Those resting on a previous wrongful act. These may be called the duties of reparation. (2) Some rest on previous acts of other men, i.e. services done by them to me. These may be loosely described as the duties of gratitude. (3) Some rest on the fact or possibility of a distribution of pleasure or happiness (or of the means thereto) which is not in accordance with the merit of the persons concerned; in such cases there arises a duty to upset or prevent such a distribution. These are the duties of justice. (4) Some rest on the mere fact that there are beings in the world whose condition we can make better in respect of virtue, or of intelligence, or of pleasure. These are the duties of beneficence. (5) Some rest on the fact that we can improve our own condition in respect of virtue or of intelligence. These are the duties of self-improvement. (6) I think that we should distinguish from (4) the duties that may be summed up under the title of ‘not injuring others’. No doubt to injure others is incidentally to fail to do them good; but it seems to me clear that non-maleficence is apprehended as a duty distinct from that of beneficence, and as a duty of a more stringent character. It will be

¹⁰ I should make it plain at this stage that I am assuming the correctness of some of our main convictions as to prima facie duties, or, more strictly, am claiming that we know them to be true. To me it seems as self-evident as anything could be, that to make a promise, for instance, is to create a moral claim on us in someone else. Many readers will perhaps say that they do not know this to be true. If so, I certainly cannot prove it to them; I can only ask them to reflect again, in the hope that they will ultimately agree that they also know it to be true. The main moral conviction of the plain man seem to me to be, not opinions which it is for philosophy to prove or disprove, but knowledge from the start; and in my own case I seem to find little difficulty in distinguishing these essential convictions from other moral convictions which I also have, which are merely fallible opinions based on an imperfect study of the working for good or evil of certain institutions or types of action.
noticed that this alone among the types duty has been stated in a negative way. An attempt might no doubt be made to state this duty, like the others, in a positive way. It might be said that it is really the duty to prevent ourselves from acting either from an inclination to harm others or from an inclination to seek our own pleasure, in doing which we should incidentally harm them. But on reflection it seems clear that the primary duty here is the duty not to harm others, this being a duty whether or not we have an inclination that if followed would lead to our harming them; and that when we have such an inclination the primary duty not to harm others gives rise to a consequential duty to resist the inclination. The recognition of this duty of non-maleficence is the first step on the way to the recognition of the duty of beneficence; and that accounts for the prominence of the commands ‘thou shalt not kill’, ‘thou shalt not commit adultery’, ‘thou shalt not steal’, ‘thou shalt not bear false witness’, in so early a code as the Decalogue. But even when we have come to recognize the duty of beneficence, it appears to me that the duty of non-maleficence is recognized as a distinct one, and as prima facie more binding. We should not in general consider it justifiable to kill one person in order to keep another alive, or to steal from one in order to give alms to another.

The essential defect of the ‘ideal utilitarian’ theory is that it ignores, or at least does not do full justice to, the highly personal character of duty. If the only duty is to produce the maximum of good, the question who is to have the good—whether it is myself, or my benefactor, or a person to whom I have made a promise to confer that good on him, or a mere fellow man to whom I stand in no such special relation—should make no difference to my having a duty to produce that good. But we are all in fact sure that it makes a vast difference. ...

It is necessary to say something by way of clearing up the relation between prima facie duties and the actual or absolute duty to do one particular act in particular circumstances. If, as almost all moralists except Kant are agreed, and as most plain men think, it is sometimes right to tell a lie or to break a promise, it must be maintained that there is a difference between prima facie duty and actual or absolute duty. When we think ourselves justified in breaking, and indeed morally obliged to break, a promise in order to relieve some one’s distress, we do not for a moment cease to recognize a prima facie duty to keep our promise ... We have to distinguish from the characteristic of being our duty that of tending to be our duty. Any act that we do contains various elements in virtue of which it falls under various categories. In virtue of being the breaking of a promise, for instance, it tends to be wrong; in virtue of being an instance of relieving distress it tends to be right.

Something should be said of the relation between our apprehension of the prima facie rightness of certain types of act and our mental attitude towards particular acts. It is proper to use the word ‘apprehension’ in the former case and not in the
latter. That an act, *qua* fulfilling a promise, or *qua* effecting a just distribution of good, or *qua* returning services rendered, or *qua* promoting the good of others, or *qua* promoting the virtue or insight of the agent, is *prima facie* right, is self-evident; not in the sense that it is evident from beginning of our lives, or as soon as we attend to the proposition for the first time, but in the sense that when we have reached sufficient mental maturity and have given sufficient attention to the proposition it is evident without any need of proof, or of evidence beyond itself. It is self-evident just as a mathematical axiom, or the validity of a form of inference, is evident. The moral order expressed in these propositions is just as much part of the fundamental nature of the universe (and, we may add, of any possible universe in which there were moral agents at all) as is the spatial or numerical structure expressed in the axioms of geometry or arithmetic. In our confidence that these propositions are true there is involved the same trust in our reason that is involved in our confidence in mathematics; and we should have no justification for trusting it in the latter sphere and distrusting it in the former. In both cases we are dealing with propositions that cannot be proved, but that just as certainly need no proof.

... Our judgements about our actual duty in concrete situations have none of the certainty that attaches to our recognition of the general principles of duty. A statement is certain, i.e. is an expression of knowledge, only in one or other of two cases: when it is either self-evident, or a valid conclusion from self-evident premises. And our judgements about our particular duties have neither of these characters. (1) They are not self-evident. Where a possible act is seen to have two characteristics, in virtue of one of which it is *prima facie* right, and in virtue of the other *prima-facie* wrong, we are (I think) well aware that we are not certain whether we ought or ought not to do it; that whether we do it or not, we are taking a moral risk. We come in the long run, after consideration, to think one duty more pressing than the other, but we do not feel certain that it is so. And though we do not always recognize that a possible act has two such characteristics, and though there may be cases in which it has not, we are never certain that any particular possible act has not, and therefore never certain that it is right, nor certain that it is wrong. For, to go no further in the analysis, it is enough to point out that any particular act will in all probability in the course of time contribute to the bringing about of good or of evil for many human beings, and thus have a *prima facie* rightness or wrongness of which we know nothing. (2) Again, our judgements about our particular duties are not logical conclusions from self-evident premises. The only possible premises would be the general principles stating their *prima facie* rightness or wrongness *qua* having the different characteristics they do have; and even if we could (as we cannot) apprehend the extent to which an act will tend on the one hand, for example, to bring about advantages for our benefactors, and on the other hand to bring about disadvantages for fellow men who are not our benefactors, there is no principle by which we can draw the conclusion that it is on the whole right or on the whole wrong. In this respect the judgement as to the rightness of a particular
act is just like the judgement as to the beauty of a particular natural object or work of art. A poem is, for instance, in respect of certain qualities beautiful and in respect of certain others not beautiful; and our judgement as to the degree of beauty it possesses on the whole is never reached by logical reasoning from the apprehension of its particular beauties or particular defects. Both in this and in the moral case we have more or less probable opinions which are not logically justified conclusions from the general principles that are recognized as self-evident.

There is therefore much truth in the description of the right act as a fortunate act. If we cannot be certain that it is right, it is our good fortune if the act we do is the right act. This consideration does not, however, make the doing of our duty a mere matter of chance. There is a parallel here between the doing of duty and the doing of what will be to our personal advantage. We never know what act will in the long run be to our advantage. Yet it is certain that we are more likely in general to secure our advantage if we estimate to the best of our ability the probable tendencies of our actions in this respect, than if we act on caprice. And similarly we are more likely to do our duty if we reflect to the best of our ability on the prima facie rightness or wrongness of various possible acts in virtue of the characteristics we perceive them to have, than if we act without reflection. With this greater likelihood we must be content.

Many people would be inclined to say that the right act for me is not that whose general nature I have been describing, viz. that which if I were omniscient I should see to be my duty, but that which on all the evidence available to me I should think to be my duty. But suppose that from the state of partial knowledge in which I think act A to be my duty, I could pass to a state of perfect knowledge in which I saw act B to be my duty, should I not say ‘act B was the right act for me to do’? I should no doubt add ‘though I am not to be blamed for doing act A’. But in adding this, am I not passing from the question ‘what is right’ to the question ‘what is morally good’? At the same time I am not making the full passage from the one notion to the other; for in order that the act should be morally good, or an act I am not to be blamed for doing, it must not merely be the act which it is reasonable for me to think my duty; it must also be done for that reason, or from some other morally good motive. Thus the conception of the right act as the act which it is reasonable for me to think my duty is an unsatisfactory compromise between the true notion of the right act and the notion of the morally good action.

The general principles of duty are obviously not self-evident from the beginning of our lives. How do they come to be so? The answer is, that they come to be self-evident to us just as mathematical axioms do. We find by experience that this couple of matches and that couple make four matches, that this couple of balls on a wire and that couple make four balls: and by reflection on these and similar discoveries we come to see that it is of the nature of two and two to make four. In
a precisely similar way, we see the *prima facie* rightness of an act which would be the fulfilment of a particular promise, and of another which would be the fulfilment of another promise, and when we have reached sufficient maturity to think in general terms, we apprehend *prima facie* rightness to belong to the nature of any fulfilment of promise. What comes first in time is the apprehension of the self-evident *prima facie* rightness of an individual act of a particular type. From this we come by reflection to apprehend the self-evident general principle of *prima facie* duty. ...

[There is an important difference between rightness and mathematical properties. A triangle which is isosceles necessarily has two of its angles equal, whatever other characteristics the triangle may have—whatever, for instance, be its area, or the size of its third angle. ... And the same is true of all mathematical attributes. It is true, I may add, of *prima facie* rightness. But no act is ever, in virtue of falling under some general description, necessarily actually right; its rightness depends on its whole nature¹¹ and not on any element in it. The reason is that no mathematical object (no figure, for instance, or angle) ever has two characteristics that tend to give it opposite resultant characteristics, while moral acts often (as everyone knows) and indeed always (as on reflection we must admit) have different characteristics that tend to make them at the same time *prima facie* right and *prima facie* wrong; there is probably no act, for instance, which does good to anyone without doing harm to someone else, and *vice versa*.

Supposing it to be agreed, as I think on reflection it must, that no one means by ‘right’ just ‘productive of the best possible consequences’, or ‘optimific’, the attributes ‘right’ and ‘optimific’ might stand in either of two kinds of relation to each other. (1) They might be so related that we could apprehend *a priori*, either immediately or deductively, that any act that is optimific is right and any act that is right is optimific, as we can apprehend that any triangle that is equilateral is equiangular and *vice versa*. Professor Moore’s view is, I think, that the coextensiveness of ‘right’ and ‘optimific’ is apprehended immediately. He rejects the possibility of any proof of it. Or (2) the two attributes might be such that the question whether they are invariably connected had to be answered by means of an inductive inquiry. Now at first sight it might seem as if the constant connection of the two attributes could be immediately apprehended. It might seem absurd to suggest that it could be right for anyone to do an act which would produce consequences less good than those which would be produced by some other act in his power. Yet a little thought will convince us that this is not absurd. The type of case in which it is easiest to see that this is so is, perhaps, that in which one has

¹¹ To avoid complicating unduly the statement of the general view I am putting forwards I have here rather overstated it. Any act is the origination of a great variety of things many of which make no difference to its rightness or wrongness. But there are always many elements in its nature (i.e. in what it is the origination of) that make a difference to its rightness or wrongness, and no element in its nature can be dismissed without consideration as indifferent.
made a promise. In such a case we all think that *prima facie* it is our duty to fulfil the promise irrespective of the precise goodness of the total consequences. And though we do not think it is necessarily our actual or absolute duty to do so, we are far from thinking that any, even the slightest, gain in the value of the total consequences will necessarily justify us in doing something else instead. Suppose, to simplify the case by abstraction, that the fulfilment of a promise to A would produce 1,000 units of good\(^{12}\) for him, but that by doing some other act I could produce 1,001 units of good for B, to whom I have made no promise, the other consequences of the two acts being of equal value; should we really think it self-evident that it was our duty to do the second act and not the first? I think not. We should, I fancy, hold that only a much greater disparity of value between the total consequences would justify us in failing to discharge our *prima facie* duty to A. After all, a promise is a promise, and is not to be treated so lightly as the theory we are examining would imply. What, exactly, a promise is, is not so easy to determine, but we are surely agreed that it constitutes a serious moral limitation to our freedom of action. To produce the 1,001 units of good for B rather than fulfil our promise to A would be to take, not perhaps our duty as philanthropists too seriously, but certainly our duty as makers of promises too lightly.

Or consider another phase of the same problem. If I have promised to confer on A a particular benefit containing 1,000 units of good, is it self-evident that if by doing some different act I could produce 1,001 units of good for A himself (the other consequences of the two acts being supposed equal in value), it would be right for me to do so? Again, I think not. Apart from my general *prima facie* duty to do A what good I can, I have another *prima facie* duty to do him the particular service I have promised to do him, and this is not to be set aside in consequence of a disparity of good of the order of 1,001 to 1,000 though a much greater disparity might justify me in so doing.

Or again, suppose that A is a very good and B a very bad man, should I then, even when I have made no promise, think it self-evidently right to produce 1,001 units of good for B rather than for A? Surely not. I should be sensible of a *prima facie* duty of justice, i.e. of producing a distribution of goods in proportion to merit, which is not outweighed by such a slight disparity in the total goods to be produced.

Such instances—and they might easily be added to—make it clear that there is no self-evident connection between the attributes 'right' and 'optimific'. The theory we are examining has a certain attractiveness when applied to our decision that a particular act is our duty (though I have tried to show that it does not agree with

---

\(^{12}\) I am assuming that good is objectively quantitative, but not that we can accurately assign an exact quantitative measure to it. Since it is of a definite amount, we can make the supposition that its amount is so-and-so, though we cannot with any confidence assert that it is.
our actual moral judgements even here). But it is not even plausible when applied to our recognition of *prima facie* duty. For if it were self-evident that the right coincides with the optimific, it should be self-evident that what is *prima facie* right is *prime facie* optimific. But whereas we are certain that keeping a promise is *prima facie* right, we are not certain that it is *prima facie* optimific (though we are perhaps certain that it is *prima facie* bonific). Our certainty that it is *prima facie* right depends not on its consequences but on its being the fulfilment of a promise. The theory we are examining involves too much difference between the evident ground of our conviction about *prima facie* duty and the alleged ground of our conviction about actual duty.

The co-extensiveness of the right and the optimific is, then, not self-evident. And I can see no way of proving it ... If, as I have tried to show, for an act to be right and to be optimific are not the same thing, and an act’s being optimific is not even the ground of its being right, then if we could ask ourselves (though the question is really unmeaning) which we ought to do, right acts because they are right or optimific acts because they are optimific, our answer must be ‘the former’. ...

In what has preceded, a good deal of use has been made of ‘what we really think’ about moral questions; a certain theory has been rejected because it does not agree with what we really think. It might be said that this is in principle wrong; that we should not be content to expound what our present moral consciousness tells us but should aim at a criticism of our existing moral consciousness in the light of theory. Now ...[i]t would be a mistake to found a natural science on ‘what we really think’, i.e. on what reasonably thoughtful and well educated people think about the subjects of the science before they have studied them scientifically. For such opinions are interpretations, and often misinterpretations, of sense-experience; and the man of science must appeal from these to sense-experience itself, which furnishes his real data. In ethics no such appeal is possible. We have no more direct way of access to the facts about rightness and goodness and about what things are right or good, than by thinking about them; the moral convictions of thoughtful and well-educated people are the data of ethics just as sense-perceptions are the data of a natural science. Just as some of the latter have to be rejected as illusory, so have some of the former; but as the latter are rejected only when they are in conflict with other more accurate sense-perceptions, the former are rejected only when they are in conflict with other convictions which stand better the test of reflection. The existing body of moral convictions of the best people is the cumulative product of the moral reflection of many generations, which has developed an extremely delicate power of appreciation of moral distinctions; and this the theorist cannot afford to treat with anything other than the greatest respect. The verdicts of the moral consciousness of the best people are the foundation on which he must build; though he must first compare them with one another and eliminate any contradictions they may contain.
A Critique of Utilitarianism
from Utilitarianism: For and Against
by Bernard Williams (1973)

Let us look...at two examples, to see what utilitarianism might say about them, what we might say about utilitarianism and, most importantly of all, what would be implied by certain ways of thinking about the situations. ...

(1) George, who has just taken his Ph.D. in chemistry, finds it extremely difficult to get a job. He is not very robust in health, which cuts down the number of jobs he might be able to do satisfactorily. His wife has to go out to work to keep them, which itself causes a great deal of strain, since they have small children and there are severe problems about looking after them. The results of this, especially on the children, are damaging. An older chemist, who knows about this situation, says that he can get George a decently paid job in a certain laboratory, which pursues research into chemical and biological warfare. George says that he cannot accept this, since he is opposed to chemical and biological warfare. The older man replies that he is not too keen on it himself, come to that, but after all George's refusal is not going to make the job or the laboratory go away; what is more, he happens to know that if George refuses the job, it will certainly go to a contemporary of George's who is not inhibited by any such scruples and is likely if appointed to push along the research with greater zeal than George would. Indeed, it is not merely concern for George and his family, but (to speak frankly and in confidence) some alarm about this other man's excess of zeal, which has led the older man to offer to use his influence to get George the job... George's wife, to whom he is deeply attached, has views (the details of which need not concern us) from which it follows that at least there is nothing particularly wrong with research into chemical and biological warfare. What should he do?

(2) Jim finds himself in the central square of a small South American town. Tied up against the wall are a row of twenty Indians, most terrified, a few defiant, in front of them several armed men in uniform. A heavy man in a sweat-stained khaki shirt turns out to be the captain in charge and, after a good deal of questioning of Jim which establishes that he got there by accident while on a botanical expedition, explains that the Indians are a random group of the inhabitants who, after recent acts of protest against the government, are just about to be killed to remind other possible protestors of the advantages of not protesting. However, since Jim is an honored visitor from another land, the captain is happy to offer him a guest's privilege of killing one of the Indians himself. If Jim accepts, then as a special mark of the occasion, the other Indians will be let off. Of course, if Jim refuses, then there is no special occasion, and Pedro here will do what he was about to do when Jim arrived, and kill them all. Jim, with some desperate recollection of schoolboy fiction, wonders whether if he got hold of a gun, he could hold the captain, Pedro and the
rest of the soldiers to threat, but it is quite clear from the set-up that nothing of that kind is going to work: any attempt at that sort of thing will mean that all the Indians will be killed, and himself. The men against the wall, and the other villagers, understand the situation, and are obviously begging him to accept. What should he do?

To these dilemmas, it seems to me that utilitarianism replies, in the first case, that George should accept the job, and in the second, that Jim should kill the Indian. Not only does utilitarianism give these answers but, if the situations are essentially as described and there are no further special factors, it regards them, it seems to me, as obviously the right answers. But many of us would certainly wonder whether, in (1), that could possibly be the right answer at all; and in the case of (2), even one who came to think that perhaps that was the answer, might well wonder whether it was obviously the answer. Nor is it just a question of the rightness or obviousness of these answers. It is also a question of what sort of considerations come into finding the answer. A feature of utilitarianism is that it cuts out a kind of consideration which for some others makes a difference to what they feel about such cases: a consideration involving the idea, as we might first and very simply put it, that each of us is specially responsible for what he does, rather than for what other people do. This is an idea closely connected with the value of integrity. It is often suspected that utilitarianism, at least in its direct forms, makes integrity as a value more or less unintelligible. I shall try to show that this suspicion is correct.

... Our descriptions of these situations have not so far taken account of how George or Jim will be after they have taken the one course or the other; and it might be said that if they take the course which seemed at first the utilitarian one, the effects on them will be in fact bad enough and extensive enough to cancel out the initial utilitarian advantages of that course. ...

There is a powerful and recognizable appeal that can be made on this point: as that a refusal by Jim to do what he has been invited to do would be a kind of self-indulgent squeamishness. ... The reason why the squeamishness appeal can be very unsettling, and one can be unnerved by the suggestion of self-indulgence in going against utilitarian considerations, is not that we are utilitarians who are uncertain what utilitarian value to attach to our moral feelings, but that we are partially at least not utilitarians, and cannot regard our moral feelings merely as objects of utilitarian value. Because our moral relation to the world is partly given by such feelings, and by a sense of what we can or cannot 'live with', to come to regard those feelings from a purely utilitarian point of view, that is to say, as happenings outside one's moral self, is to lose a sense of one's moral identity; to lose, in the most literal way, one's integrity. At this point utilitarianism alienates one from one's moral feelings ...
Suppose that there is in a certain society a racial minority. Considering merely the ordinary interests of the other citizens, as opposed to their sentiments, this minority does no particular harm; we may suppose that it does not confer any very great benefits either. Its presence is in those terms neutral or mildly beneficial. However, the other citizens have such prejudices that they find the sight of this group, even the knowledge of its presence, very disagreeable. Proposals are made for removing in some way this minority. If we assume various quite plausible things (as that programs to change the majority sentiment are likely to be protracted and ineffective) then even if the removal would be unpleasant for the minority, a utilitarian calculation might well end up favoring this step, especially if the minority were a rather small minority and the majority were very severely prejudiced, that is to say, were made very severely uncomfortable by the presence of the minority.

A utilitarian might find that conclusion embarrassing ...