



The Additive Fallacy

Author(s): Shelly Kagan

Source: *Ethics*, Vol. 99, No. 1 (Oct., 1988), pp. 5-31

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/2380927>

Accessed: 13-08-2014 16:16 UTC

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *Ethics*.

<http://www.jstor.org>

The Additive Fallacy

Shelly Kagan

I. CONTRAST ARGUMENTS

Much moral philosophy is concerned with defending or attacking the moral relevance of various distinctions. Thus consequentialists disagree with deontologists, and deontologists disagree among themselves, over whether any moral weight should be given to such distinctions as that between what one does and what one merely allows, or to the distinction between what one intends as a means, and what one merely foresees as a side-effect, and so on. Similarly, there is disagreement over the moral relevance of such factors as the motive of the agent, the consequences of a given act, or the guilt of those who may be harmed. (On some matters, perhaps, there is widespread agreement: there seems to be a contemporary consensus, e.g., that skin color is of no intrinsic moral importance.)

Such discussions, of course, are of intrinsic theoretical interest, for as moral philosophers we would like to have adequate beliefs about which factors (and which distinctions) are morally relevant.¹ They are also, furthermore, of some practical importance, for we appeal to such factors in assessing the moral status of actions in difficult and controversial cases, and in order to explain and defend our judgments. Obviously, however, such practical applications are of limited value until we have determined whether or not a given distinction actually merits being given moral weight. But how are we to settle this? How can we defend, or attack, claims involving the moral relevance of different distinctions?

A very common form of argument proceeds by offering a pair of cases that differ only in terms of the factor in question.² If we judge the

1. I take it that if all values of a given factor have the same moral importance, then the factor is of no genuine moral relevance. (For example, skin color is morally irrelevant, for differences in skin color make no moral difference.) On the other hand, for a factor of genuine moral relevance, it will be important to distinguish between the relevantly different values of the factor. The *distinctions* that make up a classificatory scheme for a given factor are morally relevant if and only if they compartmentalize the values of that factor into groups whose differences are morally relevant.

2. Some prominent and explicit examples of this sort of argument include: Michael Tooley, "Abortion and Infanticide," *Philosophy and Public Affairs* 2 (1972): 37–65, at pp. 58–60; Richard Trammell, "Saving Life and Taking Life," *Journal of Philosophy* 72 (1975): 131–37, 131–33, and "Tooley's Moral Symmetry Principle," *Philosophy and Public Affairs Ethics* 99 (October 1988): 5–31

© 1988 by The University of Chicago. All rights reserved. 0014-1704/89/9901-0008\$01.00

two cases to be morally different, it is argued, this difference must arise from the different values of the given factor³ (everything else being held constant), and so the factor is shown to be morally relevant. If we judge the cases to be similar, on the other hand, despite the variation in the given factor, this shows that the factor is not actually morally relevant after all. Let us call the pair of cases offered for comparison *contrast cases* (since the argument turns on the presence or absence of a contrast in our judgments about the cases), and let us call arguments of this sort *contrast arguments*.

By way of example, consider the distinction between what I do and what I (merely) allow. If I take out my gun and shoot you, I have brought about your death, killed you: I have done harm. In contrast, if I merely come upon you drowning, and fail to throw you the life preserver, I have not *done* any harm to you, even if you drown: I have only allowed harm. Most of us not only draw this distinction, we give it moral weight, holding that doing harm is worse than merely allowing harm. Now it is not easy to make the do/allow distinction precise, but this need not concern us here. Assuming that we *can* make the distinction, we still need to ask: is it morally relevant?

A critic of the do/allow distinction might argue as follows: "The issue of whether I brought about a given harm or merely allowed it is not itself of intrinsic moral importance. No doubt the distinction is typically correlated with other factors which do make a difference morally, and this has perhaps blinded us to the irrelevance of the do/allow distinction itself. For example, cases of doing harm are typically also cases of harming as a means, or cases where the agent has a repugnant motive, or cases where the consequences would be better if the act in question were avoided, and so on; cases of allowing harm, on the other hand, typically do not display these features. [Different critics will, of course, offer different lists of factors.] For this reason, cases of doing harm typically *are* worse than cases of allowing harm. But this does not show that the distinction

5 (1976): 305–13; James Rachels, "Active and Passive Euthanasia," *New England Journal of Medicine* 292 (January 9, 1975): 78–80; Peter Singer, *Practical Ethics* (Cambridge: Cambridge University Press, 1979), pp. 149–52 (cf. pp. 162–68 and 195–97); and Jonathan Bennett, "Morality and Consequences," in *The Tanner Lectures on Human Values*, vol. 2, ed. Sterling McMurrin (Cambridge: Cambridge University Press, 1981), pp. 45–116, at pp. 72–95. Bennett is by far the fullest presentation.

3. For some factors it will be important to distinguish several values: there may be, e.g., many relevantly different types of motive. For other factors, however, it may simply be a question of whether a certain feature is present or absent: e.g., do we have a case of causing harm, or don't we? In the latter situation, it will often be a matter of terminological convenience to speak of an always-present factor with two possible values (e.g., culpability, with the possible values of guilt or innocence), rather than of a one-valued factor (e.g., guilt) that may be either present or absent. It is also worth noting that in some cases we might treat a factor as having two values (e.g., guilt/innocence), while recognizing that in other cases a more fine-grained approach to the same factor may be appropriate (e.g., distinguishing between different levels of guilt).

between doing and allowing *itself* affects the moral status of an act. For if the distinction *were* morally relevant, it should make a difference even when all of these other factors are held constant. We can see that this isn't so, however, by considering this pair of contrast cases:

- a) Gertrude pushes Bertrand into a deep pit, hoping that the fall will kill him so that she will inherit the family fortune;
- b) Seeing that Bertrand is about to fall into a deep pit, Gertrude deliberately refrains from warning him, hoping that the fall will kill him so that she will inherit the family fortune.

These cases display the same morally relevant features (e.g., motive, outcome, etc.); they differ only with regard to the doing or allowing of harm. If the do/allow distinction were itself of genuine moral importance, there should be some difference in the moral status of the two acts. But surely Gertrude's behavior in the second case is quite as bad as in the first, despite the fact that in the first case she kills Bertrand, while in the second she merely allows him to die. Since the difference between doing and allowing harm does not make a difference in this case, we can see that the do/allow distinction is not itself of any intrinsic moral importance."

Armed with this conclusion, the critic may go on to derive moral implications concerning other cases. He might, for example, argue that our failure to make significant contributions to famine relief is roughly the moral equivalent of murder. For although we are only allowing the starving to die, given the moral irrelevance of the do/allow distinction, our failure to aid is morally as bad as if we were killing them. (Other differences, of course, e.g., our lack of ill will toward the starving, may ameliorate the badness of our act somewhat in comparison with typical cases of murder.)

There are a number of ways in which a defender of the do/allow distinction might attempt to defuse the critic's argument, and I will discuss some in the next section. But first it is important to see that contrast arguments can be used to advocate as well as to criticize a distinction. Thus the advocate of the do/allow distinction might argue as follows: "Admittedly, cases of doing harm typically differ from cases of allowing harm in terms of *several* morally relevant factors. But of course this does not show that the do/allow distinction lacks moral importance in its own right. If it *were* morally irrelevant, then in contrast cases where the only difference was whether harm was done or merely allowed, all other features being held constant, no moral difference between the cases should remain. But we can see that this is not so, by considering the following pair of cases:

- c) Ludwig sees Sylvia drowning, but since the rocks beneath the water would do extensive damage to his boat, he decides not to rescue her;
- d) Ludwig sees that his boat is about to hit Sylvia, but since avoiding her would mean steering into the rocks, which would do extensive damage to his boat, he decides not to change course.

Certainly there is something morally unacceptable about Ludwig's behavior in the first case; but surely his action is even worse in the second. Yet how could there be this moral difference in the status of the two cases⁴ if the do/allow distinction were morally irrelevant? After all, all of the other relevant features of the two cases are the same. So it must be that the do/allow distinction *is* of intrinsic moral importance, and what we have just established is that doing harm is morally more significant than merely allowing harm."

Armed with *this* conclusion, the advocate may go on to suggest that negative duties (not to do harm) are more strict than positive duties (to provide aid). Or she might use her conclusion in defense of the claim that an innocent person may not be harmed, even if this is the only way to save several other people from equivalent harms (see Sec. VI below). And so on.

This time it is the critic of the distinction who must try to disarm the contrast argument. But it should be clear that although the two sides may disagree about the use of particular instances of contrast arguments, they share a common belief in the general strategy behind such arguments. Both sides believe that with a properly constructed pair of contrast cases, the moral status of the cases will differ if and only if the do/allow distinction is morally relevant.

II. DIFFICULTIES IN USING CONTRAST ARGUMENTS

Obviously enough, additional contrast arguments concerning doing and allowing can be offered using other pairs of cases. It is equally obvious that contrast arguments can also be constructed (and have been constructed) concerning other disputed factors. It is worth noting, however, that contrast arguments can be used not only to test the relevance of various factors for the moral status of *acts*, but also to test the relevance of factors for *other* categories of moral concern, such as the moral character of the agent, or the moral value of states of affairs. (There is, as it were, more than one dependent variable in ethics.) In principle, a factor which is irrelevant for one category might well be relevant for some other category. For simplicity, however, I will limit our examples to arguments concerning the relevance of factors for the moral status of acts.

It should also be noted that in actual practice, contrast arguments are rarely spelled out as explicitly as in the examples I have given. Indeed, in many cases, we are not actually presented with a pair of contrast cases at all: the argument may simply draw our attention to the overall moral status of a single case with some salient feature; it may then conclude that the given feature has made a difference to the status of that case.

4. To be exact, of course, the question is not whether the cases themselves differ in moral status but, rather, whether the acts of the agents described in the particular cases differ in moral status. For simplicity of exposition, however, I will frequently use the imprecise formulation.

Such arguments are apt to be unintelligible unless we assume that there is an implied reference to the missing member of the contrast pair: it is taken to be so obvious that the moral status would differ in the corresponding case in which all other features are held constant except for the given factor, that there is no need to explicitly state this fact.

Contrast arguments, then, are an important and widely used method of tackling questions about the relevance of disputed factors. Both advocates and critics of the factor in question propose appropriate contrast cases, in the shared belief that examination of such cases will indicate whether or not the given factor has genuine moral weight. And since both sides accept the validity of the general strategy behind contrast arguments, each side is troubled by the existence of contrast cases that appear to support the other side. For example, as already noted, the advocate of the do/allow distinction must do something to impugn the negative argument based on cases *a* and *b*; and the critic of the distinction must do something to undermine the positive argument based on cases *c* and *d*. In such disputes, each side must attempt to disarm the contrast arguments offered by the other side.

There are several different approaches that have been commonly used to argue that—despite initial appearances—a given pair of contrast cases fails to support the desired conclusion concerning the relevance of some factor. This is because although the strategy of contrast arguments is generally accepted, it is not always an easy matter to apply the strategy properly. Here are some of the things that can go wrong:

i) Since contrast arguments are based on the idea that the effect (or lack of effect) of a given factor will be exposed when all other factors are held constant, a pair of contrast cases is useful only if all other things really *are* held constant. If some other factor is varied asymmetrically in the pair of cases (e.g., if the motive is worse in one case than in another), then the differential impact of this other factor may well mask the effect (or lack of effect) of the original factor being tested. (Critics, confronted with positive arguments, may then be able to claim that the difference in the moral status of the two cases is due entirely to the difference in the second factor. Advocates, confronted with negative arguments, may be able to claim that the similar overall moral status of the two cases is the net result of the two genuine factors cancelling out each other's individual effects.) Most criticisms of contrast arguments involve the charge that other things have not been kept equal in the contrast cases, that is, that some relevant factor (other than the one being examined) differs from one case to the other.

One might well wonder whether it is ever possible to apply the strategy adequately, for one might hold that we never *can* keep all the other morally relevant factors constant. (Perhaps differences in the factor we want to test are unavoidably correlated with differences in some other relevant factor.) This is an important worry, but I want to put it aside, noting only that this objection still seems to concede the theoretical validity

of the contrast strategy: perhaps we cannot actually keep everything else equal, but if we *could* . . .

ii) Contrast arguments are based on the idea that when all factors are held constant except for the one being examined, if the factor in question is of genuine moral relevance, this will make a difference to the moral status of the contrast cases. Thus a pair of cases will only be of use if the two really do differ in terms of the factor being tested. Obviously, a contrast pair will not reveal anything about the importance of, for example, motive, or consequences, if both cases display the *same* motive, or the same consequences, and so on. Often it will be uncontroversial whether this condition has been satisfied, but not always. For example, a critic of the do/allow distinction might offer a pair of cases that differ only in terms of whether exotic life-support equipment for a terminal patient is turned off or merely allowed to malfunction, arguing that the moral status of these cases is the same. Such a contrast argument might be attacked by a defender of the distinction on the grounds that turning off life-support equipment in such a case is merely allowing the patient to die and is not actually a case of doing harm at all.

Criticisms of this sort may force us to specify, much more exactly than we have previously done, criteria for determining when we have instances of one or the other side of the given distinction (or of specific values of the factor in question, when more than two are possible, as in, e.g., motive). Such specifications may well be controversial, so there will remain room to criticize contrast arguments that rely on them.

iii) Contrast arguments turn on whether a pair of properly constructed cases differ in moral status. Inevitably, however, such arguments can only appeal to our intuitive judgments about the given cases. Since few would hold that our intuitions in such matters are infallibly accurate or completely precise, it is possible to question whether our intuition is sufficiently sensitive to detect moral differences in all cases in which they are present. In some cases, for example, it might be suggested that one or another of the factors that are being held constant is so forceful that the moral status of both cases is extreme—so extreme, in fact, that the remaining difference between the cases is relatively slight in comparison. In such situations, the difference between the cases—although genuine—is lost on our insufficiently sensitive intuitions. Thus both cases may be so horrible, or so wonderful (or even so obviously acceptable) that our failure to notice any difference in the moral status of the cases may not be accepted as sufficient evidence that such a difference is actually lacking. Such a challenge can be met by trying to construct contrast cases where the factors being held constant are “toned down,” so that any possible difference resulting from the variation in the factor under examination will stand out better.

Even if the other factors are subdued, so that the impact of the given factor will have a better chance of being detected, it still might be the case that although the factor makes a genuine difference it is such a

slight one that our intuition remains insufficiently sensitive to detect it. However, although critics of the given factor must admit that this remains a possibility, even the advocate of the factor will have to concede that if this is indeed the situation then the factor is of negligible importance, and for practical purposes can be overlooked.

iv) Since a contrast argument has persuasive force only for those who share the intuitive moral assessment of the contrast cases, it remains open for someone to reject a given argument on the grounds that he sincerely fails to make the judgments that the argument presupposes. It is in order to avoid such challenges, of course, that those who offer contrast arguments strive to find cases where the intuitive judgments are uncontroversial. Focusing on such uncontroversial cases has two advantages. First, this obviously allows the argument to retain potential persuasive force against some who might otherwise escape it. Second, since the presence of disagreements in intuitive judgments for a particular case calls into question, at the very least, the confidence that we can place in the *accuracy* of our intuition concerning that case, focusing on uncontroversial cases allows the argument to proceed where we have the greatest confidence in our intuitive judgments.

This last point, however, raises a much more radical objection: perhaps for some range of cases we should not trust our intuitions at all. It might be held that our intuitive moral judgments may well be wildly inaccurate (even when there is a consensus), telling us that one act is morally worse than another, when in fact the opposite is the case. Such skepticism about the accuracy of our intuitions might be held because our intuition is thought to be morally corrupt, or because it is thought that there is simply no reason to give our intuitions any particular credence, or for a variety of other reasons. Such worries should, I think, be taken more seriously by all who are not intuitionists, but once again I note the problem only to put it aside: *most* contemporary moral philosophy relies heavily on appeals to moral intuitions, and for better or for worse, contrast arguments are no different in this regard.

These, then, are some of the basic ways that particular contrast arguments have been criticized. Such arguments turn on comparing the moral status of well-constructed contrast cases. As such, they have been attacked on the ground that the cases described are not actually well constructed, whether because the factor under examination does not vary properly between the two cases (ii), or because all other factors are not held constant (i). Similarly, such arguments have been attacked on the ground that our assessment of the contrast cases is impaired, whether because of a lack of sufficient precision in our intuition (iii), or because of a more general lack of dependability and accuracy (iv). I want to emphasize, however, that although all of these criticisms can be used to attack particular contrast arguments, they share a theoretical acceptance of the general strategy behind such arguments: if we *were* to construct an adequate pair of contrast cases, and if we *did* accurately assess the

moral status of those cases, then a moral difference in the cases would appear if and only if the given factor were of intrinsic moral relevance.

III. THE UBIQUITY THESIS

The account of contrast arguments I have just given normally lies mostly implicit in our use of such arguments; but I believe most will recognize it as a reasonable portrayal of the thinking that lies behind our common practice. I have spent a fair amount of space developing this account in order to draw attention both to the wide use of such arguments and to their plausibility. Contrast arguments are based on a clear and plausible strategy, and that strategy can be accompanied by a well-articulated set of considerations that must be met if the strategy is to be properly applied.

Nonetheless, I believe that our in-principle acceptance of contrast arguments must be mistaken. Despite its plausibility, the strategy relies on an underlying assumption concerning the role of factors—an assumption that is questionable and should probably be rejected. Yet without this assumption, contrast arguments cannot be used to derive the conclusions we draw from them. (A much more limited use for such arguments may remain; see Sec. VII below.)

The contrast strategy clearly assumes that if a factor has genuine moral relevance, then for *any* pair of cases, where the given factor varies while others are held constant, the cases in that pair will differ in moral status. (The difference, of course, need not be so extreme as to make, e.g., one act permissible and the other act impermissible; the point is simply that *some* difference in moral status should exist.) That is, putting aside difficulties with proper construction and assessment, in principle *any one pair of cases* should be sufficient to settle the question of whether the given factor is of intrinsic moral relevance. In effect, the contrast strategy must be assuming that if variation in a given factor makes a difference *anywhere*, it makes a difference *everywhere*. Let us call this the *ubiquity thesis*.

Contrast arguments presuppose the ubiquity thesis. Without this assumption, after all, from the mere fact that some particular pair of contrast cases did not differ in moral status, critics of the given factor could not go on to conclude that the factor must lack intrinsic moral significance—for it might well make a difference in other cases, even though it did not make one here. Similarly, without this assumption, from the mere fact that some particular pair of contrast cases *did* differ in moral status, advocates of the given factor could not go on to conclude that the factor must make a difference in other cases where it is present—for it might well fail to do so someplace else even though it did make a difference here. Indeed, recognizing that the contrast argument makes this assumption—that if the factor makes a difference anywhere it must make one everywhere—helps to explain why both critics and advocates are at such pains to disarm the contrast cases offered by the other side: if the assumption is correct, a single well-constructed and properly assessed pair of cases from either side must be decisive.

The ubiquity thesis is in need of some defense. Why must a factor make a difference everywhere if it makes one anywhere? There is a strong temptation to think that this must follow (indeed, follow trivially) from the fact that the fundamental principles of morality are universal,⁵ that is, that the basic principles that determine the moral status of an act do not vary from case to case. Admittedly, the applicability of derivative principles may change with the situation, but the underlying principles of morality do not. But this means that the role that a given genuine moral factor plays in determining the moral status of an act must be universal as well: it too cannot vary from case to case. Therefore—or so it seems to follow—if a factor makes a difference *somewhere* it must make a difference *everywhere*.

Despite the seductiveness of this line of thought, I believe it is mistaken. I do find it plausible to think that the universal nature of fundamental moral principles implies that the role a genuine factor plays must be universal as well. But I do not think that we can infer the ubiquity thesis from this. That thesis asserts that if a factor ever makes a difference to the overall moral status of an act, it must always do so. That is, it asserts that a particular kind of *effect* must be universal. Obviously, however, the universality of that particular effect will only follow from the universality of the role that the factor plays, if it is part of that role that it have that effect. Once again, it may seem trivial to assert that every effect that a factor has is part of its role (or, at the very least, that this particular effect must be part of its role); but in fact this depends on substantive claims about the roles of genuine moral factors.

An analogy may be helpful here. The presence or absence of oxygen has a role in determining chemical reactions. This role is presumably universal—that is, the fundamental laws of chemistry do not vary from case to case. Yet, obviously enough, the particular effects of oxygen's presence or absence do vary: in some cases, for example, the presence of oxygen makes a difference to whether or not a compound burns; but it would be a mistake to think that it must make this difference in every case. This is because making a difference to whether compounds burn is not, strictly speaking, part of the role of oxygen in the laws of chemistry. Rather, it is a consequence of that role in particular cases. The mere fact that the role of oxygen in the laws of chemistry is universal does not imply that some particular kind of effect must be universal. Similarly, the mere fact that the role of some factor in the fundamental principles of morality must be universal does not automatically imply that some particular effect must be universal. Whether the ubiquity thesis is true depends on taking a substantive position on the role of genuine moral factors.

5. Bennett is one of the few who explicitly recognizes that the contrast argument assumes the ubiquity thesis. He suggests that the assumption follows from the universal nature of basic reasons (pp. 73–74; cf. p. 92). Unfortunately, Bennett does little to explain this suggestion, but it may be that he is making what I will go on to call the “additive assumption.”

Talk of the role of moral factors seems to naturally suggest the following general model: the moral status of an act is a function of the various (genuinely relevant) factors that come into play in the given situation. Different moral theories will, of course, hold different views about which factors are of intrinsic moral importance; but whatever the particular list, all theories agree that the overall status of an act in a given situation depends on the values that the genuine factors (whatever they are) have in that situation. This agreement, however, is limited to the view that there is a function of *some* sort connecting the values of the factors to the overall status of the act. Exactly *how* the status of the act is thought to be affected by the factors depends on what *particular* function is accepted by the given moral theory.

Let us call the function (whatever it is) that is taken to determine the overall status of the act on the basis of the values of the factors the *governing function*. In principle, even theories that agree about the list of relevant factors might disagree about what the governing function is like, and thus disagree about how the factors combine and interact in determining the act's overall moral status.

When we say that the role of a given moral factor is universal, we are saying that the governing function remains the same from case to case. That is, the rules governing the interaction of the various moral factors in determining the moral status of an act remain constant.

But this does not yield the ubiquity thesis. Depending on the particular governing function accepted by a theory, it might well be that variation in a given factor would make a difference to the moral status of the act in certain situations, but not in others—because of differences in the interactions of the various moral factors arising from the differences in the values of the *other* factors. Only for particular kinds of governing functions will the ubiquity thesis hold true. What I want to suggest is that those who offer contrast arguments presuppose a particular view of the nature of the governing function.

IV. THE ADDITIVE ASSUMPTION

The view which underlies the use of contrast arguments seems to be this: the function that determines the overall status of the act given the values of the particular factors is an *additive* one. That is, the status of the act is the net balance or sum which is the result of adding up the separate positive and negative effects of the individual factors. On this view, each factor makes a contribution, whether positive or negative, to the moral status of the act. (The strength of the contribution will depend on the particular value of the factor in the given situation, e.g., how *much* good will be done by the act.) The overall status is the sum of these positive and negative contributions.

In effect, each of the positive factors provides a reason *for* performing the act (the strength of the reason depending on the value of the factor), and each of the negative factors provides a reason for *not* performing the act. The moral status of the act depends on whether the combined

weight of the reasons for performing the act is greater or less than the combined weight of the reasons for not performing the act—and by how much.

It is worth stressing that the view that the governing function is additive is more than the mere suggestion that the overall status is the sum of two amounts—that is, the joint positive contribution and the joint negative contribution. Rather, each of these amounts is in turn taken to be the sum of the individual contributions of the individual factors. Each factor makes its own contribution to the status of the act, and the overall status is the sum of these separate contributions.

It is certainly not obvious that the contrast argument assumes that the governing function is additive. Yet I think that something like this assumption must be made for there to be any reason to believe that a well-constructed pair of contrast cases will differ in moral status if and only if the given factor has intrinsic moral relevance.

It is easy to see why one will believe in the ubiquity thesis—and thus in the strategy behind contrast arguments—if one believes that the governing function is additive. For if the overall status of an act is the sum of the separate contributions made by the individual factors, then if one holds all but one of those contributions fixed (by holding all but one of the factors constant) and varies the remaining contribution (by varying the factor under examination), the sum will have to vary as well. The mathematical analogy should make this clear. If $S = x + y + z$, then if we hold x and y constant, but vary z , S will have to vary as well. *Any* genuine difference in z will have to make a difference to S , no matter *what* the values of x and y —provided only that x and y are kept constant. (The same is true, of course, for each of the other variables.) Similarly, then, if the governing moral function is additive, if we hold all but one factor constant, no matter what contributions those particular factors make, varying the remaining factor will have to vary the remaining contribution, and thus will have to make a difference to the sum.

Now for some purported factors, no doubt, variations in the factor will make no difference to the contribution the factor makes. But this is just to say that the factor is not of any genuine moral relevance. Since it will contribute equally to the moral status of all acts, it can just as easily—and less misleadingly—be dropped from consideration altogether. (Consider, e.g., the claim that being performed in time makes a positive contribution to the moral status of the act, a contribution whose size is constant, no matter what the time of performance.) For any genuine factor, variations will affect the contribution, and thus the sum.⁶ And as

6. There might also be cases where variations in a genuine factor did not actually result in a variation in the contribution that the factor makes. But this would be tantamount to showing that the variation under consideration was morally irrelevant. (One might believe, e.g., that doing harm is worse than merely allowing it, and nonetheless hold that certain differences in how harm is done—e.g., with a gun or with a knife—are of no moral relevance.) To be precise, no doubt, one should speak of morally relevant variations in morally relevant factors, but for obvious reasons I have avoided this cumbersome locution.

we have just seen, variation will affect the sum no matter what the contributions of the other factors. Thus for a genuine factor, differences in that factor will always make a difference to the overall moral status of the act. But this, of course, is the ubiquity thesis.

Thus the assumption that the governing function is additive—the *additive assumption*—is sufficient to ground the strategy behind contrast arguments. Furthermore, the additive assumption seems plausible in its own right. The model of morally relevant factors making positive and negative contributions to the overall moral status of the act is a familiar and attractive one. As I have noted, the thought that typically there are reasons for and against an act naturally lends itself to the view that the governing function is additive. I suggest, therefore, that those who offer contrast arguments have been tacitly (or explicitly) presupposing the additive assumption.

It should be noted, however, that the ubiquity thesis does not itself entail the additive assumption. That is, there are logically possible governing functions which satisfy the ubiquity thesis although they are not themselves additive. Therefore, commitment to contrast arguments does not itself logically presuppose acceptance of the additive assumption.⁷ However, despite this logical possibility, it is difficult to think of any other independently attractive views about the nature of the governing function which reject the additive assumption and yet nonetheless manage to provide plausible reason to believe that the ubiquity thesis will still be satisfied. Therefore I want to stick to my conjecture that contrast arguments are accepted because—and only because—of the belief that the governing function is additive.

I have, however, been somewhat inaccurate in my presentation of the additive assumption. To be more precise, I have smuggled in a second assumption concerning the nature of the governing function. In discussing the implications of the additive assumption, I have done more than assume that the overall status of the act is the sum of the separate contributions of the individual factors. I have also presupposed that the size of a given factor's contribution is determined solely by the value of that factor: variations in the given factor will affect the nature of its contribution; but variations in the *other* factors will *not* affect the contribution made by

7. The ubiquity thesis actually assumes that for each genuine factor, whenever all other factors are kept constant, the governing function will be equal to a function of the given factor which is one-to-one. That is, with the values of other factors fixed, no two (relevantly different) values of the given factor will be mapped onto (i.e., determine) the same overall moral status. The assumption that the governing function is "one-to-one" in this way is weaker than the additive assumption. It is also worth noting in this regard that mere satisfaction of the ubiquity thesis does not itself guarantee that a given value of the factor will always make a contribution of the same size (nor even that the direction of the contribution will be constant—i.e., always positive or always negative). Contrast arguments do not require these stronger conditions, although many believe that they are satisfied as well. (See the discussion of transport arguments in Sec. VI; and cf. n. 11 below.) My thanks go to Peter Vallentyne and Ken Manders for discussion of these matters.

the factor in question. That is, I have assumed that the size of a factor's contribution is *independent* of everything other than the value of the factor itself.

Strictly speaking, the assumption that the governing function is additive does not entail that the contributions of the individual factors are independent. One might believe that the overall status of an act is the sum of the contributions made by individual factors, while at the same time holding that the size of the contribution of a given factor is determined in part by the values of other factors. Note that with such a view there would be no particular reason to assume that the ubiquity thesis was true, for it might well be the case that although variation in a given factor would make a difference to its contribution in most situations, in *certain* situations (because of the particular configuration of the other factors) such variation might make *no* difference to the contribution, and thus none to the overall sum. It is only by assuming that the contribution a factor makes is independent of the other factors that there is some reason to think that such a possibility is avoided.

It should also be noted that the assumption of independent contributions does not entail the additive assumption. One might hold, for example, that the overall moral status of an act is not the result of adding but rather of *multiplying* some of the separate, independent contributions of the individual factors—perhaps adding the rest. Once again, however, on such a view there would be no particular reason to assume that the ubiquity thesis was true.

The mathematical analogy should make this clear. If $S = x \cdot y + z$, then although differences in y would normally make a difference to S when x and z are held constant, in those cases where x is equal to zero, differences in y would *not* affect S . Similarly, if the contributions of some of the moral factors are multiplied, then it might well be in certain cases that even variation in a genuine factor would make no difference to the overall moral status of the act. (No doubt it would be somewhat strained to speak of the factor's *contribution* in such a case.) It is only by assuming that the governing function is additive that there is any plausible reason to assume that this possibility is avoided.

Thus it would be more precise to say that those who offer contrast arguments appear to be presupposing both that the governing function is additive and that individual contributions are independent. Despite the logical separability of these two assumptions, however, they combine naturally in the view I have been describing. The very term "contribution" strongly suggests both assumptions. It would seem somewhat strained to speak of the contribution of an individual factor, if in fact that contribution was determined not only by the given factor but by the other factors as well. And as already noted, talk of contributions seems somewhat out of place in contexts which are not additive.

All of this suggests that it might be more accurate to speak of contrast arguments as presupposing "the model of independent, summed con-

tributions”—or perhaps “the contribution assumption.” Nonetheless, for purely mnemonic reasons (and possibly idiosyncratic ones, at that) I am going to continue referring to “the additive assumption”—hereafter understanding it to include independence, and not just additiveness *per se*. My reason is simply that I find this label effective in quickly conjuring up the entire model I have in mind, while other labels are less effective in this regard.

My claim, then, is that one should accept the strategy behind contrast arguments only if one accepts the additive assumption. It is not so much that contrast arguments *could* not be sound if the additive assumption is false, but rather that there is no plausible reason to think that they *are* sound if the additive assumption is false.

V. THE PLAUSIBILITY OF THE ADDITIVE ASSUMPTION

The view that the moral status of an act is the sum of individual positive and negative contributions—the particular reasons for and against performing the act—is, I have suggested, a familiar and attractive one. Nonetheless, I believe that the additive assumption should be rejected. More exactly, I believe that the assumption is a controversial one, likely to be false on most moral theories. To assume the truth of the additive model without defense, therefore, is unacceptable: it begs the question against most plausible theories, and it may well be incompatible with one's own considered views.

One or two examples should be sufficient to show that most theorists will need to reject the additive assumption. Consider, first, the issue of self-defense. Most of us believe that in some situations killing someone in self-defense can be morally permissible. Imagine such a situation, and compare these two cases:

- e) In order to defend myself against the aggressor, I push him into a pit, expecting the fall to kill him;
- f) In order to defend myself against the aggressor, I refrain from warning him about the pit into which he is about to fall, and I expect the fall to kill him.

Most of us will certainly want to claim that the moral status of these two cases is the same, even though, obviously enough, in the first case I do harm, while in the second case I merely allow it. Not everyone will accept this judgment: some may insist that there is a slight difference, but it is lost on our insufficiently sensitive intuitions; others (e.g., pacifists) may claim the difference is rather significant. Nonetheless, most of us will want to maintain that the distinction between doing harm and allowing harm makes *no* difference in such legitimate cases of self-defense. The crucial point to see is that even the *advocate* of the do/allow distinction may well want to make this claim, and surely it is coherent for her to do so. In fact, this view—that although the difference between doing harm and allowing it typically has a great effect on the moral status of an act,

in cases of self-defense it makes no difference whatsoever—is not only coherent: most of us want to claim that it is true.

Yet if the additive assumption is correct, then the view I have just described *must* be false, since it violates the ubiquity thesis to claim that although the do/allow distinction normally makes a difference, in cases of self-defense it does not. Thus, if we make the additive assumption (thereby justifying the ubiquity thesis), this view can be ruled out a priori. Obviously, however, such a move would simply beg the question. On the other hand, if we want at the very least to allow for the possibility of such views—we cannot simply presuppose that the governing function is additive. The example of self-defense suggests that most of us (unwittingly, perhaps) are committed to the rejection of the additive assumption.

It seems a mistake to treat the factor of self-defense as though it made an independent contribution to the overall moral status of the act—a contribution to be added to that made by other factors. A more natural interpretation, one closer to our intuitive understanding of the situation, would be to view the factor of self-defense as acting more like a zero multiplier.

Recall the earlier mathematical example where $S = x \cdot y + z$, and suppose for simplicity that x can equal only zero or one. When x is equal to one, differences in y will make a difference to S ; but when x is equal to zero, differences in y will not matter at all. The self-defense/non-self-defense factor seems to act somewhat like x in this example: normally, with a value of one, it allows the do/allow distinction to make a difference; but in cases of self-defense, it takes on the value of zero, and so differences between doing and allowing harm do not affect the moral status of the act at all.

I would not want to put too much weight on this mathematical analogy; I certainly do not think it captures all that we intuitively want to say about the interaction of self-defense and the do/allow distinction. But it clearly comes closer than viewing the governing function in strictly additive terms. It seems reasonable to suspect that intuitively adequate representations of most views in this area would have to stray even further from the additive path. This does not prove, of course, that the correct governing function is nonadditive. But it does, I think, show that most will want to accept this conclusion, and that arguments that presuppose the opposite will tend to beg the question.

A second example suggests that even in some areas where we may initially seem able to give an additive formulation, we may nonetheless require a nonadditive one if we are to capture more of what we want to say. Most of us believe that the alleviation of suffering is a morally relevant factor: the fact that an act will relieve suffering provides a reason for performing that act (although the reason can, of course, be outweighed by other considerations). Thus, when all things are equal, if we can aid only one of two people in pain, there is greater reason to help the one

whose suffering is greater. Suppose, however, that all things are not equal, and a second factor comes into play: imagine that the person whose pain is greater—Trixie—is partially responsible for the plight of both, while the other person—Fritz—is not responsible at all. Most of us want to say of such situations that there is somewhat less reason to aid Trixie than there would be were she free from responsibility; and if her responsibility is significant enough, we may judge it better to aid Fritz, even though he is in considerably less pain.

Now we could, I suppose, view this in an additive manner. The magnitude of Trixie's suffering makes an independent, positive contribution to the moral status of aiding her. This is offset by her partial responsibility, which makes an independent negative contribution—that is, provides a reason for not aiding her. The sum of these two contributions may well be less than the corresponding sum for aiding Fritz, where the admittedly smaller positive contribution is not offset by any negative factor.

Such a model may seem plausible so long as we imagine Trixie morally blameless, despite her partial responsibility. The problem is that the additive assumption requires this same model even when Trixie is morally to blame for having performed the act which created the situation. Perhaps, for example, Trixie had attempted to harm Fritz, but her scheme had gone awry, harming them both, and her more significantly. The additive model can point to her guilt as a reason for not helping Trixie, but it is forced to maintain that the magnitude of her suffering provides just as strong a reason to aid her as if she were innocent. Yet many would want to reject exactly this latter claim: the suffering of the guilty simply does not count as much as the suffering of the innocent.

Obviously, not all will share this judgment; but those who do will find it plausible to reject the additive assumption. It will seem more appropriate to view the factor of innocence/guilt not as making an independent contribution, but rather as a multiplier—this one able to range from one to zero. Fritz is innocent, so the multiplier has the value of one, and his suffering can take on its full potential weight. Trixie is guilty, so the multiplier is less than one, and her suffering counts less. Potentially, with significant enough guilt, the multiplier would be equal to zero, and so one's suffering would not count at all. (A more pronounced version of this view might allow the multiplier to range between positive one and negative one. If the wicked deserve to be unhappy, then the presence of suffering and desert should together sometimes yield a positive contribution.)

Once again, I certainly would not want to claim that this alternative model captures all that we want to say about the interplay of desert and well-being. But it seems a more natural fit than the additive model,⁸ and

8. This alternative model gives up the belief that contributions of the individual factors must be summed, but it continues to view those contributions as separate and independent. It might be suggested that it would be more natural still to give up the independence

it suggests that those sympathetic to such views should reject the additive assumption.

One possible response to the examples I have given is that they merely show the difficulty of capturing certain moral views in an additive model using the *particular* factors I have described. Some might suggest, in the light of this, that the additive assumption could be preserved in a given case provided that we appeal instead to a different, more adequate list of the morally relevant factors. This possibility should certainly be conceded. Indeed, there is a sense in which it is trivially true.⁹ The question, however, is whether one can describe an acceptable list of morally relevant factors that can be incorporated into an additive model across the board: moral views that escape one or two particular examples may fall prey to others. Since each example relies on substantive claims about moral theory, no one example will convince everyone that the additive assumption needs to be rejected. But I hope it is clear that examples such as the two I have given could easily be multiplied.

Reflection on such examples, I believe, should support two conclusions. First, since a great number of moral views are incompatible with the additive assumption, it is unacceptable to criticize such views through the use of arguments that simply presuppose that assumption: to do so merely begs the question. Second, since one's own moral views are likely to be incompatible with the additive assumption, it is dangerous to defend those views through the use of arguments that presuppose that assumption: to do so may render one's position incoherent.

I suspect, in fact, that for most views the additive model is grossly inadequate. It is not merely that the claim that the *entire* governing function is additive is false. Rather, on most views, *most* of the governing function will fail to be additive. For most factors, their role in determining the overall moral status of an act simply cannot be adequately captured in terms of separate and independent contributions that merely need to be added in.

If this is right—if the implausibility of the additive model is so easily seen—the question becomes: what draws us to the additive assumption

condition. One could avoid the need to talk of multiplying the separate contributions of the two factors, by holding that there is a single contribution here—that of the suffering—but that the nature of this contribution is not independent of the factor of desert. It is not altogether clear to me what turns on this disagreement, but what is important for our purposes is that this model, too, would entail rejecting the additive assumption.

9. Provided that one is willing to accept unnatural and ad hoc factors, there will always be a way to save the additive assumption. After all, one could insist that there is only *one* genuine factor: the complete specification of the given act's place in the history of the world. The additive assumption would hold trivially on this view: the contribution made by this global factor would satisfy independence (for there are no other factors on which to be dependent); and the overall status of the act would be the "sum" of this single contribution. The interesting question is whether there is a satisfying way to specify factors below the global level and still maintain the additive assumption. For most views, I doubt it.

in the first place? What explains the persistent temptation to assume an additive model? "Persistent temptation" seems exactly the right phrase here. In my own case, at least, I find myself naturally gravitating toward the additive assumption; despite my conscious intention to avoid simply assuming that factors should be treated in an additive fashion, I catch myself unwittingly making just this assumption nonetheless.

Part of the attraction may lie in the fact that the additive assumption seems a natural generalization from a plausible treatment of simple cases. Suppose that a young child is drowning, and I can save him at negligible cost (and there are no other morally relevant features to the case). There is an almost irresistible account of this case, which fairly seems to stare us in the face: the fact that the child can be saved from death provides a *reason* for taking steps to save him—a reason of a certain strength.

Imagine, next, that in order to rescue the child, I would have to tell a lie (e.g., to the owner of the boat). Intuitively, this provides a reason for *not* taking the necessary course of action—a reason of a certain strength. But this does not affect the fact that there is still a reason *for* taking the necessary steps—the fact that it would save the child's life. The correct account seems obvious: there is a reason for and a reason against, and the permissibility of the act depends upon which reason is stronger. Furthermore, the strength of each individual reason seems independent of the presence or absence of the other reason. The need to lie provides a reason not to perform the act, which may or may not be outweighed on balance; but the strength of that reason in itself does not seem affected by the amount of good that will be done by the lie. Similarly, the reason for taking the boat—that it saves a life—may or may not be outweighed on balance; but it does not seem in itself any weaker than in the original case where it was the only relevant reason.

In short, the additive model suggests itself as the obvious account for the simple case I have just described: the two factors make their own independent contributions, and the status of the act is determined by adding—balancing—these separate contributions. And it is natural to assume that the extra complications caused by bringing in other factors can be handled in the same manner. For example, if the drowning child is *my* child, this seems to give me an extra reason for taking the boat—a reason to be added to the others. And so on.

Thinking about cases like these may tempt us into the general conclusion that the governing function is additive. Especially when the question is never consciously posed, such cases make it easy to tacitly assume that the additive model is correct. But the temptation should be resisted. It may well be that in isolated cases it is illuminating to think of the morally relevant factors as making individual positive or negative contributions to the moral status of the act; but there is no reason to assume that this model has any general applicability. As I have suggested, for most moral theories the interplay of moral factors is simply too complicated to be adequately captured in additive terms.

Yet the additive assumption remains attractive nonetheless. In fact, the very drive to understand moral judgments can pull us toward it. After all, it is hardly philosophically helpful to view such judgments as black boxes, the grounds of which cannot be analyzed. And if we want to understand the moral status of an act, no analytic tool could be more basic than the simple question: what reasons are there to perform the act, and what reasons are there not to perform it? But now the very language of reasons makes it easy to think in additive terms. If there are various reasons for and against an act, then moral judgment is a matter of evaluating and weighing the individual reasons, and balancing them. Indeed, the image of a balance scale immediately suggests itself: each reason falls like a coin with a certain weight into one or the other of the two pans in the scale; the greater the combined weight in a given pan, the more it tips the balance in its own favor. And since each reason must be grounded in some morally relevant feature of the situation, progress in understanding is gained by isolating the particular factor that generates the given individual reason.

Thus we arrive, once again, at the additive assumption. The view is a natural one to hold, and at times it feels almost irresistible. As we have seen, however, the additive model is questionable enough that there is certainly no reason to assume that it is true without argument. Indeed, on many views, there is actually reason to believe that it is false.¹⁰

VI. TRANSPORT ARGUMENTS

Given the seductiveness of the additive assumption, it is not surprising that most of us rely on it frequently and unwittingly. Once we make the assumption, some extremely attractive forms of argument become available to us. Throughout this paper, I have, of course, stressed the reliance of contrast arguments on the additive assumption. But an even more powerful—and quite common—form of argument relies on this assumption as well. I will call these arguments *transport arguments*.

Most cases are complex enough that, even if we can agree on the list of morally relevant factors, it is difficult to know how to proceed so as to arrive at and defend a judgment concerning the overall moral status of the act in question. But if the additive assumption is correct, a method naturally suggests itself, for we can isolate the individual contributions of particular factors. By constructing simple enough cases, we can isolate particular features and note the strength of the reason that a given

10. G. E. Moore's defense, in *Principia Ethica*, of an organic theory of the good can be seen as a rejection of the additive assumption in a particular area of value theory. Indeed, my entire discussion of the additive assumption can be viewed, with hindsight, as a generalization of Moore's point to all of ethics, together with an application of this generalization to areas where its relevance has not been recognized. Something similar to the additive assumption has also been criticized by Michael Philips, in an article published after I had completed this essay ("Weighing Moral Reasons," *Mind* 96 [1987]: 367–76).

feature grounds. We can then “transport” this information back into complex and more controversial cases: observing the presence of the same feature, we can infer that it grounds a reason of the same strength. Coupled with similar knowledge concerning the other factors in the complex case, we can then justify a judgment about the overall status of the act by summing the individual positive and negative contributions.¹¹

Actual applications of this strategy are, of course, subject to various practical difficulties, and this gives room for criticism of particular arguments. But in many cases, it should be noted, even fairly rough assessments of the individual contributions (or merely comparative orderings of their relative strengths) may be enough to support a given judgment.

Provided we are making the additive assumption, the strategy behind such transport arguments seems legitimate. For so long as we view factors as making separate and independent contributions, there is reason to think that when we “transport” a given feature from situation to situation it will continue to ground a reason of the same strength. Recall the image of coins in a balance: each individual coin always makes the same particular contribution, no matter what other coins are in the pans as well.

If the additive assumption is false, however, then the strategy behind transport arguments is illegitimate, for the “contribution” of a factor will not be fixed in this way. For example, if the contribution is dependent not only on the single feature, but on other factors as well, then despite the preservation of the feature considered in isolation, a changed context could easily result in a significantly altered contribution. Similarly, even if the contributions are independent, but are not simply to be summed, then a feature which in one context combines with the other factors so as to ground a reason for performing an act might, in other contexts, combine with the altered factors to ground a reason of a different strength, or even a reason which *opposes* the given act.

Thus, if the additive assumption is false, the sort of argument that straightforwardly combines the results concerning the impact of factors on simple cases into conclusions about more complex cases is called into question. (The same is true, of course, for arguments that transport results from particular complex cases to *different* complex cases.) For such arguments typically turn on attempts to isolate the separate contributions of individual factors, and without the additive assumption such attempts are ill-conceived.

Given the ease with which one can slip into making the additive assumption unwittingly, however, it is often difficult to recognize that a given transport argument is relying on a controversial and potentially question-begging premise. The assumption is rarely articulated, and even

11. Contrast arguments, it will be recalled, assume that if a given factor makes a difference somewhere, it will make a difference everywhere; but they do not necessarily assume that it will make the *same* difference everywhere (see n. 7 above). Transport arguments, however, clearly make this bolder assumption as well. Obviously, if the bolder assumption is true, the more modest one will also be true.

when one realizes the need to make it, it is readily agreed to. It may be helpful, therefore, to consider one or two examples carefully.

Do Numbers Matter?

Imagine that I have to choose between saving one innocent person, who will otherwise die, or saving five other innocent people, who will otherwise die. Which group should I save, the one or the five? Assuming there are no other morally relevant factors, almost all of us would say it is better to save the larger group. Given that—unfortunately—I cannot manage to get to both groups in time to save all six, it is better that five should live than one: there is greater reason to save the five than there is to save the one.

Not everyone, however, believes that the numbers matter in this way. Given that someone will be saved no matter which group I choose, these people claim that there is no more reason to save the five than the one.¹² Since it would be improper simply to dismiss this “anti-number” position dogmatically, various arguments have been offered in defense of the more common “pro-number” position.

One argument asks us to consider, first, a simple case where we must decide whether or not to save Claudia (as opposed to doing nothing). Obviously, if all other things are equal, there is a significant reason to save her: this is surely what we mean when we say that morally her life counts for something. Next, consider a different case where we must choose whether to save Albert or to save Bernard, for we cannot manage to save both. Here, if all other things are equal, there is surely no more reason to pick the one than the other: we are faced with a tie. Finally, consider the case where we must choose between saving only Albert, or saving both Bernard and Claudia. Since the choice between Albert and Bernard was a tie, and the opportunity to save Claudia’s life counts for *something*, the tie must now be broken: there is greater reason to save the two than there is to save only the one. (If $X = Y$, and $Z > 0$, then $X + Z > Y$.) Thus the numbers do count after all, and the anti-number position must be mistaken.¹³

For years, I found this argument persuasive. But now it seems clear to me that it simply presupposes the additive assumption, and thus begs the question. Admittedly, the possibility of saving Claudia’s life is a morally relevant factor. But it is only by illegitimately making the additive as-

12. This position has been defended by John M. Taurek, “Should the Numbers Count?” *Philosophy and Public Affairs* 6 (1977): 293–316; and, I believe, by G. E. M. Anscombe, “Who Is Wronged?” *Oxford Review*, vol. 5 (1967).

13. I learned this argument from T. M. Scanlon (who, however, does not endorse it). More complicated versions of the argument can be found in Jonathan Glover, *Causing Death and Saving Lives* (New York: Penguin, 1977), pp. 207–9, and in Gregory S. Kavka, “The Numbers Should Count,” *Philosophical Studies* 36 (1979): 285–94, pp. 291–92; but I believe that both of these make mistakes similar to the ones I discuss below. Both Glover and Kavka offer additional arguments against the anti-number position, as does Derek Parfit in “Innumerate Ethics,” *Philosophy and Public Affairs* 7 (1978): 285–301.

sumption that from the argument it can be inferred that this feature makes an independent contribution that can simply be added to the contributions of the other factors, without regard to the changed context.

The advocate of the anti-number position readily admits that the possibility of saving a life—rather than doing nothing—grounds a reason of a certain strength. But without the additive assumption, little can be inferred from this. For example, it does not follow that the possibility of saving two lives—rather than doing nothing—grounds a reason twice as strong. And even if this latter conclusion were accepted as well, it would not follow that there was more reason to save the two than the one. For without the additive assumption, there are no grounds for assuming that the reason generated by the possibility of saving two lives is as strong when the alternative is not that of merely doing *nothing*, but rather that of doing something of genuine moral *value*, namely, saving the one. Thus even if there is more reason to save two (rather than doing nothing) than to save one (rather than doing nothing), it does not follow that there is more reason to save two rather than saving one.

Since the advocate of the anti-number position denies that the relevance of human lives can be correctly captured in an additive manner, it is obvious that arguments against that position must not simply presuppose the additive model. It is because it is so easy to make the additive assumption without realizing it, that arguments that actually beg the question can be thought persuasive. None of this, of course, shows that the anti-number position is correct. It may well be that the factor of human lives *should* be treated in an additive manner; but it will take a stronger argument to prove this.

Do/Allow Again

For a second example, let us return to the controversy over the importance of the do/allow distinction. Imagine that the only way I can save the lives of Gustav and Emile is by killing Philippa, and suppose, for simplicity, that there are no other morally relevant factors. Despite the fact that my killing Philippa would result in there being two alive rather than only one, most of us would say that it is wrong to kill her. Many would appeal to the moral relevance of the do/allow distinction in defense of this judgment, noting that I face a choice between *doing* harm to Philippa and merely *allowing* harm to Gustav and Emile.

However, even if we grant the relevance of the distinction between doing and allowing harm, it is not obvious exactly how this justifies the prohibition against killing Philippa. To clarify this view, someone defending the prohibition might argue as follows:

“Consider a case in which I can save two lives, by donating \$1,000. While it would be meritorious of me to make this sacrifice, I am not morally required to do so. However, since I *would* have to save the two lives if I could do so at negligible cost, presumably the size of the sacrifice is a morally relevant factor; in this case, it evidently grounds a reason

for my not making the sacrifice of sufficient strength to counter the reason *for* making it (i.e., that it would save two lives). Now consider the different case in which I can save \$1,000 by killing someone. This is clearly impermissible, despite the fact that the same size sacrifice is at stake as in the first case, and there are even fewer lives involved. If killing the one to save \$1,000 is forbidden, while allowing two to die rather than spending \$1,000 is permitted, this must be because the reason for not killing one is stronger than the reason for not allowing two to die. But this means that if the only way to save the lives of Gustav and Emile is by killing Philippa, there is greater reason to refrain from killing her. For killing Philippa is only supported by the reason for not allowing the other two to die, while refraining is supported by the reason for not killing someone—and we have just seen that the latter reason is stronger than the former reason.”

Having worked through the earlier example, it should now be clear that despite the plausibility of this argument, it too relies on the additive assumption. Even if we grant the intuitive assessment of the two initial cases, the conclusion does not follow unless we assume an additive model. First of all, the argument clearly assumes that the overall moral status of each of these cases is the sum of two contributions: on the one hand, a reason in favor of killing/letting die, based on the cost to me of doing otherwise; and on the other hand, a reason which opposes that act, either on the grounds that it is a killing or on the grounds that it is a letting die. Second, it assumes that the strength of the reason based on the cost to me is the same in both cases, presumably on the grounds that the strength of this reason is determined solely by the size of the sacrifice. Without both of these assumptions, there is no reason to infer that the reason for not killing one is stronger than the reason for not allowing two to die. Yet without the additive assumption, there is no particular reason to believe either.

Furthermore, even if we accept the argument’s conclusion so far, it still does not follow that there is greater reason not to kill Philippa in our original case. For without the additive assumption, there is no particular justification for assuming that the reason not to kill has the same strength, regardless of whether the reason for killing is merely that it will save \$1,000 or that it will save two innocent lives. Thus, even if the reason for not killing some one person (when the only reason for doing so is that it will save \$1,000) is stronger than the reason for not allowing two people to die (when the only reason for doing so is that it will save \$1,000), it still does not follow that there is greater reason not to kill Philippa when the reason for doing so is that it will save the lives of both Gustav and Emile.

Had the argument under examination been successful, it would have been part of an illuminating account of why there is a prohibition against harming some so as to prevent harm to others. Since it relies on the dubious additive assumption, however, I believe the argument must be

rejected—at least in the absence of a defense of that assumption. It is obvious, of course, that the rejection of this argument does not itself prove that the prohibition in question cannot be adequately supported. But still, it does deprive that prohibition of an intuitively plausible and attractive defense.

The two examples we have considered only begin to give one a sense of the variety of transport arguments that can be found. The use of such arguments is extremely widespread in moral philosophy, and the rejection of the underlying strategy calls into question many powerful and plausible arguments. Yet once we have abandoned the additive assumption, I can see no justification for accepting the general strategy behind such arguments. It remains a possibility, of course, that in some individual cases the application of the strategy will not lead to any false conclusions. But obviously enough, in the absence of a justification for the general strategy, arguments that simply rely on that strategy will have to be rejected as unsound.

Once we have recognized the illegitimacy of making the additive assumption, therefore, we can see the need to reconsider numerous arguments that we previously found persuasive. In this paper I have drawn attention to the way in which both contrast arguments and transport arguments rely on this assumption. The use of such arguments, I believe, is quite common. But I suspect that reliance on the additive assumption is even more widespread than this suggests, and that further forms of argument need to be questioned as well. Given the naturalness of making the additive assumption, it should not surprise us to find that it has crept unrecognized into many of the arguments offered in moral philosophy. Nor should it surprise us to find that identifying these arguments will often be difficult: many such arguments, after all, are quite silent about their (tacit) reliance on an assumption to which we readily—albeit illegitimately—agree.

VII. DOING WITHOUT THE ADDITIVE ASSUMPTION

As we have seen, the additive assumption naturally supports a project of trying to identify the separate and independent contributions of individual factors. When we reject the additive assumption, however, there is no particular reason to believe that this project can meet with any general success. At best, one might instead attempt to identify *clusters* of features which together always combine to generate a reason of a particular strength for or against performing an act. On such an alternative approach, there would be no telling in advance how many factors would be involved in any given cluster (and perhaps in rare cases a “cluster” might actually be composed of a single factor), but once one had identified such a cluster, its contribution could be added to those of the other clusters present in a given case. Admittedly, one would still be weighing and balancing *reasons* on this approach, but a model of this sort would recognize that reasons are not the contributions of individual factors taken in isolation.

Such an approach does get beyond the simplicity of the additive assumption, but it reveals its seductive influence nonetheless. For this view still thinks in terms of separate contributions (of individual clusters, now, rather than individual factors) that are added in determining the overall moral status of the act. Now in some cases it may be possible to separate out individual clusters of factors, each of which makes its own independent contribution. But it should be stressed that, in the absence of an argument, there is no particular reason to assume that even this is possible.

Well, at the very least, can't we assume that the features of a case can be divided into two groups—that is, those which jointly ground reasons *for* performing the act, and those which jointly ground reasons for *not* performing the act? This seems modest enough, but even this can only be agreed to with care. For it may be that the strength of the case for performing the act is not independent of the factors that intuitively make up the case for *not* performing the act, and vice versa. (Put another way: we must not assume that membership in the two groups will be exclusive.) Indeed, in some cases it simply may not be illuminating to attempt to sort the features into groups of positive and negative factors.

I do not know if this last, gloomy possibility is ever actually realized. But I suspect that it is the lingering influence of the additive assumption that creates the mild temptation simply to dismiss it out of hand.

Abandoning the additive assumption obviously does not mean that it will no longer be possible to offer arguments concerning the relevance and particular roles of individual moral factors. But it does, I think, mean that it will be much more difficult to offer satisfactory arguments in this area. For example, as we have seen, without the additive assumption contrast arguments simply cannot be used to support the general conclusions that are normally derived from them. Nonetheless, even without that assumption, I believe that one may be able to derive certain modest conclusions from arguments of this sort.

Imagine a pair of contrast cases that differ in overall moral status, apparently because of variation in some particular factor. Armed with the additive assumption, an advocate of that factor could have gone on to derive forceful conclusions about the importance of that factor in *other* cases. Having given up the additive assumption, however, those conclusions cannot be so readily demonstrated. Nonetheless, a more modest conclusion does still seem in order. Assuming that the contrast cases are well-constructed and properly assessed, if they differ in moral status it *does* seem to show that the factor in question is of genuine moral relevance. For surely *something* has to account for the difference in status of the two cases, and by hypothesis all other factors have been ruled out.

At this point, however, a pessimistic position suggests itself: even if it is true that a given factor can thus be shown to be morally relevant in some particular case, nothing of any general theoretical importance can be inferred from this. For without the additive assumption, we cannot

conclude that the factor will make a similar difference in other cases; we can only conclude that it would do so in exactly similar cases. But this conclusion is so restricted in scope as to be of no use for theoretical purposes.

Such pessimism seems too hasty, however, for a somewhat more moderate conclusion may be possible. Even without the additive assumption, a demonstration that a factor does make a difference in one case may still create a *presumption* that it does so in other cases as well. After all, if the factor has a particular effect in one case, there will have to be some *reason* why it does not have the same effect in other cases. This shifts the burden of proof: those who wish to deny that the factor has the given effect in some particular case must point to an attending difference in some second feature (or group of features) and offer a plausible account of how that second feature interacts with the given factor so as to alter its effects. Similarly, if a factor is shown to make no difference to a particular pair of contrast cases, one cannot infer that it makes no difference anywhere at all. But such a demonstration does create a presumption against that factor: those who claim that it does matter elsewhere will need to point to some second feature and offer a plausible account of why it affects the impact of the original factor.¹⁴

It would be equally mistaken, however, to replace our hasty pessimism with a hasty optimism. To believe that the additive assumption should be rejected is to believe that factors frequently do interact in such a way as to alter the particular effects of a given factor from case to case. Thus we should expect that it will be quite common for the presumptions I have just described to be appropriately overridden. But this means that we will have to enter into the detailed discussions of the roles of particular factors after all. There is no shortcut.

What we are forced to do is to start producing plausible theories of the interactions of the various morally relevant factors. And if such theories are not to be ad hoc, presumably they will have to be based on theoretical accounts of what makes the individual factors relevant in the first place.

14. It might be suggested, however, that there is an important asymmetry in the power of positive and negative contrast arguments. For a successful positive argument can establish that the factor in question is of genuine moral relevance, and must appear in the governing function (even if the factor will not have a noticeable effect anywhere else). A negative argument, on the other hand, cannot establish that the given factor is *not* of genuine moral relevance and does *not* appear in the governing function—for the factor might still make a difference someplace else. I believe, however, that this apparent asymmetry is not a deep one. A positive argument tells us that variation in a particular factor makes a difference in a specific context. A negative argument tells us that variation in a particular factor makes *no* difference in a specific context. The information gained is symmetrical. I suspect that the appearance of asymmetry is actually an artifact of our preference for having our theories stated only in terms of what is of moral *relevance*, rather than what is of moral *irrelevance*.

Mere examination of cases is not likely to be of much help in the production of such accounts.

Ultimately, then, the rejection of the additive assumption will force us to put less emphasis on the intuitive assessment of cases, and more emphasis on the construction of fundamental moral theory. But this, I believe, is all to the good.¹⁵

15. I owe a debt to Dan Brock, who offered the original contrast argument that set me wondering what was wrong with them.