

The inefficacy objection to consequentialism and the problem with the expected consequences response

Mark Bryant Budolfson¹

© Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract Collective action problems lie behind many core issues in ethics and social philosophy—for example, whether an individual is required to vote, whether it is wrong to consume products that are produced in morally objectionable ways, and many others. In these cases, it matters greatly what we together do, but yet a single individual’s ‘non-cooperative’ choice seems to make no difference to the outcome and also seems to involve no violation of anyone’s rights. Here it is argued that—contrary to influential arguments by Peter Singer, Alastair Norcross, Shelly Kagan, Derek Parfit, and Allan Gibbard—an appeal to the expected consequences of acts cannot deliver plausible verdicts on many of these cases, because individuals often have a probability of making a difference that is sufficiently small to ensure that ‘non-cooperation’ is the option with the greatest expected value, even when consequentialists themselves agree that ‘cooperation’ is required. In addition, an influential argument by Singer, Norcross, and Kagan is shown to be unsound for the claim that in the collective action situations at issue, the expected effect of one individual’s action equals the average effect of everyone’s similar actions. These results have general implications for normative theory, because they undermine the sort of consequentialist explanation of collective action cases that is initially attractive from many theoretical points of view, consequentialist and otherwise.

Keywords Consequentialism · Normative ethics · Peter Singer · Alastair Norcross · Shelly Kagan · Derek Parfit · Allan Gibbard · Inefficacy · Vegetarianism · Consumer ethics · Ethics of consumption · Voting

✉ Mark Bryant Budolfson
budolfson@princeton.edu

¹ Woodrow Wilson School and University Center for Human Values, Princeton University, Princeton, NJ, USA

A single individual often feels powerless in the face of the collective action of others. From a philosophical point of view, such feelings of inefficacy are often unjustified. After all, even if a single act of charity cannot solve the world's problems, it may make all the difference to a single recipient who would otherwise have died, in which case it is simply untrue that such an act *makes no difference*. And from a deontological perspective, even if it makes no difference to the outcome whether, for example, a single individual joins in a group assault, nonetheless joining would violate the victims' rights in a straightforward way, and so it is simply untrue that *there is no decisive reason* not to perform such an act. As these examples indicate, inefficacy is often an illusion, or at least is no excuse for bad behavior that directly violates rights.

However, it is far from clear that worries about inefficacy are *always* misguided, because even on the assumption that they are misguided in the kinds of cases just described, they are not obviously misguided in connection with many other core examples in ethics and social philosophy involving collective action problems—for example, whether an individual is required to vote, whether it is wrong to consume products that are produced in morally objectionable ways, and many other examples. In these cases, it matters greatly what we together do, but yet a 'non-cooperative' choice made by a single individual seems to make no difference to the welfare of others and also seems to involve no violation of rights. As a result, it is worth considering a more sophisticated inefficacy objection directed at this more limited but still crucially important range of cases, because such an objection is supported by reasoning that is not obviously at odds with the empirical facts and does not presuppose the falsity of either consequentialism or rights-based deontology.

For a concrete example of a case where the inefficacy objection is not obviously misguided—and an example that is much-discussed in the recent literature—consider an individual's decision whether to purchase and eat a factory-farmed chicken. According to utilitarians like Peter Singer, purchasing and eating such a chicken is wrong because it has unacceptable consequences on balance for welfare. However, even if we agree with Singer's premises about the magnitude of animal suffering and the comparative unimportance of gustatory and other human pleasures, the inefficacy objection is that his conclusion about the welfare effects of consumption by an individual does not follow, and, upon careful reflection, turns out to be false. That is because, according to the inefficacy objection, an individual's decision to consume animal products cannot really be expected to have any effect on the number of animals that suffer or the extent of that suffering, given the actual nature of the supply chain that stands in between individual consumption decisions and production decisions; at the same time, an individual's decision to consume animal products does have a positive effect for that individual. As a result, according to the inefficacy objection Singer's premises about animal suffering and human pleasures, together with the actual empirical facts about the workings of the marketplace, entail that the (expected) welfare effect of an individual's decision to consume animal products is *positive* on balance, in contrast to what Singer assumes. If this inefficacy objection is sound, it undermines the idea that individuals have welfare-based reasons not to eat meat from factory farms, and shows that act

utilitarian reasoning may well support conclusions that are the opposite of what utilitarians themselves endorse in connection with many analogous collective action situations in modern society.¹

Even if one rejects act utilitarianism and other forms of act consequentialism, the inefficacy objection remains important, because on most deontological views the welfare effects of one's actions constitute genuine, albeit *defeasible* ethical reasons for action, and so even on deontological views an appeal to (expected) welfare effects might seem like a promising way to explain what individuals have most ethical reason to do in collective action cases where 'non-cooperative' acts are intuitively objectionable but involve no violation of rights. Furthermore, if the inefficacy objection is sound, it also seems to show that these objectionable acts do not involve the kind of *complicity* in evil that may initially seem to provide another plausible explanation of their objectionability from a deontological perspective. For example, if, as the inefficacy objection claims, an individual's consumption of a chicken cannot be expected to make a difference to the number of chickens that suffer, then that consumption also cannot be expected to *benefit those who cause such suffering*—because then a single individual's consumption does not make a difference to the revenues of *factory farms* for the same reasons it does not make a difference to the number of animals produced on such farms. Thus, the inefficacy objection threatens to undermine the common claim that by purchasing morally objectionable products one is complicit in evil in an impermissible way because one thereby *supports* objectionable firms by *voting with one's dollars* in a way that benefits those firms.² For these reasons, the inefficacy objection is an important challenge even for deontological theories, as I argue at greater length elsewhere.³

Setting aside non-consequentialist theories for the purposes of this paper, it is clear that the inefficacy objection applies with particular force to *traditional act consequentialism* (hereafter simply *consequentialism*).⁴ That is because

¹ A useful way of focusing attention on one more general philosophical issue in connection with the inefficacy objection is by imagining a scenario in which factory farmers produce more and more animals and increase waste in the supply chains, thereby undermining any possible effect that a single individual's consumption could have—perhaps, it could be imagined, to undermine the ethical objections to their products in a utilitarian society. If factory farmers adopted such a strategy, then even if the empirical assumptions of the inefficacy objection were mistaken in the actual world, nonetheless factory farmers could in this scenario succeed in making it unobjectionable on act utilitarian grounds to consume their products—here by simply by intentionally inflicting even more suffering and acting in a way that is even more evil than their actual behavior. If this implication of act utilitarianism is implausible, it highlights the philosophical challenge of explaining what individuals have reasons to do under conditions of inefficacy. For these reasons, at least this philosophical issue raised by the inefficacy objection is independent of the actual empirical facts.

² For example, Tom Regan claims that "Since [animal agribusiness] routinely violates the rights of these animals...it is wrong to purchase its products" (Regan 2004, p. 351). Tristram McPherson's view in McPherson (2015) is similar to Regan's but is more clearly and fully developed, and in particular provides a proposal for how to analyze the relevant notion of complicity in/benefiting from wrongdoing.

³ See Budolfson (unpublished a).

⁴ Here I assume an intuitive distinction between *traditional consequentialism* and other views such as rights-based deontology that could (arguably) be expressed using the conceptual structure of consequentialism, but that would thereby be forced to endorse a theory of the good that is disunified, or at least of a kind never encountered in consequentialist theories until the late twentieth century, such as

consequentialism is intended to gain its plausibility from the idea that the consequences of acts are the only things that matter from a moral point of view. As a result, in contrast to familiar external critiques of consequentialism such as the purported counterexample of the Transplant Case,⁵ the inefficacy objection in connection with the examples described above gives rise to a powerful internal critique of the view, as consequentialists such as Peter Singer, Alastair Norcross, and Shelly Kagan have noted. Here is Kagan:

In cases of this sort, therefore, consequentialism seems to fail even by its own lights. ... Apparently then, consequentialism fails to handle a kind of case that even consequentialists admit it ought to be able to handle.⁶

It is then perhaps no surprise that consequentialists have unified behind a particular response to the inefficacy objection, which claims that it invariably overlooks the significance of low-probability threshold effects. Responses in this neighborhood have been defended by Allan Gibbard, Derek Parfit, Peter Singer, Alastair Norcross, and Shelly Kagan. Of particular note, the latter three philosophers have offered essentially the same detailed argument that such a response must be correct.⁷ Here is Norcross's presentation:

Footnote 4 continued

theories of the good that imply that there is something especially bad about performing particular types of acts *oneself*, etc. For discussion of the possibility of 'consequentializing' views such as Kantian ethics in this way, see Dreier (2011) and Portmore (2011).

⁵ For a presentation of the Transplant Case, see Thomson (1985).

⁶ Kagan (2011, p. 108).

⁷ See the passage quoted immediately below from Norcross (2004, pp. 232–233). Compare also Singer (1980, pp. 335–336): "Perhaps for every 10,000 vegetarians there is one fewer 20,000 bird chicken unit than there would otherwise be. Perhaps not: this is merely an example and I have no idea what the true figure would be; but there must be some point at which the number of vegetarians makes a difference to the size of the poultry industry. There must be a series of thresholds, hidden by the market system of distribution, which determine how many factory farms will be in existence. In this case one more person becoming a vegetarian will make no difference at all, unless that individual, added to the others who are already vegetarians, reduces demand below the threshold level at which a new factory farm would have started up (or an existing one would have remained in production, if the industry is declining). Looking at one's own decision to be a vegetarian, it may seem frustrating that one cannot be sure that one has saved even a single animal from a miserable life on a factory farm; but from a utilitarian perspective it really makes no difference whether each vegetarian is personally responsible for saving ten chickens a year from this fate, or one vegetarian in 10,000 makes the difference that will save 100,000 birds. Utilitarianism judges actions by their likely consequences, and so it ranks the certainty of saving ten chickens equally with the 1 in 10,000 chance of saving 100,000. As long as I have no idea whether or not my own decision to go vegetarian is the decision that takes the demand for chickens below the threshold, the strength of this reason for being a vegetarian is unaffected". Compare also Kagan (2011, p. 124): "...we know that there is some triggering number, T (more or less), such that every Tth purchase (more or less) triggers the order of another T chickens (more or less). I don't have any idea what that number is, but I do know that whatever it is, I have a 1 in T chance (more or less) of triggering the suffering of another T chickens (more or less). And so in terms of chicken suffering, my act of purchasing a chicken still has an expected disutility equivalent to one chicken's suffering. And since, by hypothesis, this is greater than the pleasure I will get from eating the chicken, the net expected utility of my purchase remains negative. As I walk to the butcher counter, then, not only don't I know whether my act will have bad results, I don't even know what the chances are that my act is a triggering act. But I do know, for all that, that the net expected results of my act are bad. So I should not buy a chicken". Compare also Gibbard (1990 [1971],

Suppose that there are 250 million chicken eaters in the US, and that each one consumes, on average, 25 chickens per year... Clearly, if only one of those chicken eaters gave up eating chicken, the industry would not respond. Equally clearly, if they all gave up eating chicken, billions of chickens (approximately 6.25 billion per year) would not be bred, tortured, and killed. But there must also be some number of consumers, far short of 250 million, whose renunciation of chicken would cause the industry to reduce the number of chickens bred in factory farms. The industry may not be able to respond to each individual's behavior, but it must respond to the behavior of fairly large numbers. Suppose that the industry is sensitive to a reduction in demand for chicken equivalent to 10,000 people becoming vegetarians. (This seems like a reasonable guess, but I have no idea what the actual numbers are, nor is it important.) For each group of 10,000 who give up chicken, a quarter of a million fewer chickens are bred per year. It appears, then, that if you give up eating chicken, you have only a one in ten thousand chance of making any difference to the lives of chickens, unless it is certain that fewer than 10,000 people will ever give up eating chicken, in which case you have no chance. Isn't a one in ten thousand chance small enough to render your continued consumption of chicken blameless? Not at all. While the chance that your behavior is harmful may be small, the harm that is risked is enormous.⁸

As this presentation of what might be called the Singer/Norcross/Kagan response illustrates, its main goal is to argue that in the cases at issue, for any single individual, the *expected effect* of that individual's action—in this case, purchasing one chicken—is equal to the *average effect* of all actual acts of that type, in the sense of 'expected effect' familiar from the notion of 'expectation' in expected utility theory⁹—and similarly for other collective action situations in which the inefficacy objection might initially seem to show that consequentialism has absurd implications by its own lights. If this response were correct, then consequentialism

Footnote 7 continued

pp. 26–27): "I do not accept that in cases of diffuse benefits, act-utilitarianism prescribes non-cooperation... [For example,] the net value of what n gas cheats accomplish is the sum of the values of n effects individual gas cheats could have. It is the sum of the net benefits from one gas cheat in a world with no other, the net benefit from one gas cheat in a world with two others, and so on up to a world with $n-1$ others. If the effect of n gas cheats is calamitous, at least one of these net benefits from an individual gas cheat must be negative. Hence it is possible for an individual to produce a bad result by helping to strain the gas system, no matter how uncertain and diffuse that result may be. If the system is likely to be under strain even with everyone cooperating, an act-utilitarian will cooperate. He will calculate the average expectable net benefit from an act of gas-cheating by dividing the likely effect of a large number of gas cheats by n ". Compare also Brandt (1959, pp. 389–390). Compare also Parfit (1984, pp. 73–75), who offers analogous reasoning about voting, but does not offer any clear guidance as to the intended generalizability of the reasoning.

⁸ Norcross (2004, pp. 232–233). See also Singer (1980, pp. 325–337), and Kagan (2011, p. 124), both quoted in the preceding footnote.

⁹ In particular, the 'expected effect' of an action in this sense is the *expectation* associated with that action, based on the sum of: all of the values of all of the possible outcomes of that action weighted by their probability conditional on that action.

would be able to respond to the inefficacy objection in a way that appears fully satisfactory by its own lights. Perhaps the most striking aspect of this response is that it purports to refute the inefficacy objection by means of a simple a priori demonstration that in the collective action situations at issue the expected effect of one individual's action equals the average effect of all similar actions, regardless of how the empirical facts turn out that at first glance might seem crucially important.

An initial quibble with Norcross's presentation of this response is that it seems to begin by explicitly acknowledging that the expected effect of one individual's action is zero, before going on to argue for the opposite conclusion. In particular, it begins by acknowledging that "Clearly, if only one of those chicken eaters gave up eating chicken, the industry would not respond", which appears to be merely a way of saying: "The expected effect of one individual giving up chicken is zero". For our purposes, we can simply ignore this and examine the subsequent argument.

The more serious problem with the Singer/Norcross/Kagan Response is that, as stated, it depends on invalid reasoning, and once the source of this invalidity is made clear, this undermines the idea that there could be any successful similar a priori response to the inefficacy objection, and ultimately suggests that inefficacy objection succeeds in showing that consequentialism often fails by its own lights. To begin to see why, consider the following case:

Richard makes paper T-shirts in his basement that say 'HOORAY FOR CONSEQUENTIALISM!', which he then sells online. The T-shirts are incredibly cheap to produce and very profitable to sell and Richard doesn't care about waste per se, and so he produces far more T-shirts than he is likely to need each month, and then sells the excess at a nearly break-even amount at the end of each month to his hippie neighbor, who burns them in his wood-burning stove.¹⁰ For many years Richard has always sold between 14,000 and 16,000 T-shirts each month, and he's always printed 20,000 T-shirts at the beginning of each month. Nonetheless, there is a *conceivable* increase in sales that would cause him to produce more T-shirts—in particular, if he sells over 18,000 this month, he'll produce 25,000 T-shirts at the beginning of next month; otherwise he'll produce 20,000 like he always does. So, the system is genuinely sensitive to a precise tipping point—in particular, the difference between 18,000 purchases and the 'magic number' of 18,001.

Suppose that a consumer knows all of these facts about Richard's business, and is considering buying a T-shirt for himself. What is the expected effect on the number of T-shirts produced of that consumer purchasing a T-shirt? The correct answer is essentially zero, because given what is known about the history of demand for Richard's T-shirts and how production quantities are determined, there is virtually no chance that exactly 18,001 people are going to buy Richard's T-shirts this month and trigger a dramatic threshold effect—which, of course, is not to claim that there

¹⁰ If necessary, suppose (to make the case straightforward) that because Richard is very busy, he only produces the paper T-shirts at the beginning of each month, and because the T-shirts are made out of very thin paper, the entire stock disintegrates after one month, and so there is no inventory carry-over from month to month, etc.

is *zero* chance of that happening, but rather that the odds of that happening—of exactly 18,001 of Richard’s T-shirts being sold—is certainly dramatically lower than 1/5,000 or any other number that would drive the expected effect of an individual buying one T-shirt anywhere near the consequence that 1 additional T-shirt is produced. This shows that the reasoning behind the Singer/Norcross/Kagan Response is invalid, because insofar as that response is taken to show that consuming meat should be expected to have significant bad effects for animal welfare (e.g., equal to the average effect of those individual actions), similar reasoning would show that buying one T-shirt in the story above should be expected to result in approximately 1 additional T-shirt being produced, which is the wrong result.¹¹ The problem with the reasoning is that it overlooks the fact that we can know enough about the supply chains in both cases to know that threshold effects are not sufficiently likely and are not of sufficient magnitude to drive the expected effect of consumption anywhere close to the average effect.

In response, it might be insisted that there is a crucial disanalogy between the T-shirt case just described and e.g., our actual situation with respect to factory farmed animal products. One obvious difference might seem to be the amount of waste: in particular, in the T-shirt case, a significant amount of the product is ‘wasted’. However, although the T-shirt case is indeed a dramatization aimed primarily at making vivid why the reasoning behind the Singer/Norcross/Kagan Response is invalid as presented, upon further reflection there is less of a difference to a typical case involving the consumption of animal products than it might initially appear, and more importantly there is no crucial disanalogy with respect to the expected effect of an individual’s consumption decisions. For example, consider the meat that goes out of date in a wholesaler’s meat locker or on a supermarket shelf, and is then sold to a dog food plant or ‘rendered’ into feed for other animals. Is that meat ‘wasted’? What is the difference between that meat and the ‘wasted’ T-shirts in the story above that Richard sells to his neighbor? For current purposes it doesn’t really matter whether such things are labeled as ‘wasted’ or not—what matters is that there is reliably at least a small amount of ‘wasted’-like meat at each stage of the supply chain that serves as sufficient ‘slack’ to create buffers that prevent an individual’s decision to purchase meat from making any difference to the number of animals that are produced at the far other end of the supply chain.

Most importantly, the crucial issue is not about the *magnitude* of these buffers, but rather about their *reliability*: as long as we can know—as we can—that there are sure to be buffers of non-trivial size throughout the supply chain (even if they are not nearly as large as the buffers in the T-shirt example), that reduces the probability of a single individual making a difference to a level that quickly becomes nearly

¹¹ Note that this result is not undermined by the observation (often enthusiastically made by consequentialists) that in collective action situations, as the probability of making a difference goes down, the magnitude of the difference that would be made goes up. As the discussion above illustrates, what matters is whether the difference that would be made increases in a way that is relevantly proportional to the decrease in the probability of making a difference—and as the discussion above illustrates, in real world collective action problems it is often empirically unrealistic to think that it does.

infinitesimal, in a way that is analogous to the way that mathematical models explain why the probability of casting a decisive vote in a large-scale election quickly becomes nearly infinitesimal when it is assumed that individuals have very reliable information that one candidate has a non-trivial lead over the others.¹²

Furthermore, even in the very unlikely event that, say, an individual purchase of meat really did succeed in making the price of animals at one point at a production end of the supply chain \$0.01 higher than it otherwise would have been, that would not make the dramatic difference to the number of animals that are brought into existence that it would have to make in order for the possibility of such a threshold effect to drive the expected effect toward the average effect, in part because the number of animals that are brought into existence is surprisingly insensitive to very small changes in price at that location for a variety of reasons.¹³ As a result, such a change would not have anywhere near the effect it would need to have on the number of animals produced in order to give rise to a threshold effect that would justify equating the expected effect with anything like the average effect of animal consumption decisions.

¹² For a model of this kind, see Brennan and Lomasky (1993, chapter 4). There is some controversy about whether such a model correctly represents the probability of casting a decisive vote in an election, but part of that controversy arises from the fact that the voters arguably do not have reliable enough information in voting case for such a model to be applicable (namely, information about the anticipated difference in the number of votes cast between the candidates)—but that is not a problem when discussing actual large-scale marketplaces, where investigation of their actual workings typically reveals that inefficiency, noise, etc., will give rise buffers with the high degree of reliability that the current point assumes.

¹³ To illustrate this, and to illustrate the more general point about reliable amounts of noise, inefficiency, and other drivers of individual inefficacy in actual supply chains, consider the supply chain for American beef. When ranchers who own their own grazing land decide how many cattle to raise, their decisions are sensitive to their own financial situation, the number of cattle their land can support, the expected price of any additional feed that will be needed, bull semen and other ‘raw materials’ that go into cattle production, and the expected price that the cattle will fetch when they are ultimately sold to feedlots. Of these, small changes in last—the price that cattle will fetch at the feedlot—are of the least importance, because insofar as ranchers judge that capital should be invested in raising cattle rather than other investments, they will tend to raise as many cattle as they can afford to breed and feed within that budget, letting the ultimate extent of their profits fall where it may at the feedlot. Many ranchers also use the nutritional well-being of their herd as a buffer to absorb adverse changes in market conditions, feeding their cattle less and less to whatever point maximizes the new expectation of profits as adverse conditions develop, or even sending the entire herd to premature slaughter if, say, feed prices rise to levels that are unacceptably high. This serves to shift the ranchers’ emphasis in decision making relevant to herd size even further away from the price of beef. As a result, even if an individual’s consumption decisions somehow (implausibly) managed to have a \$0.01 effect on the price of cattle at feedlots, the effect on the number of cattle produced would be much smaller than it would have to be in order for the possibility of such a threshold effect to justify equating the expected effect of an individual’s consumption of beef with the average effect of such consumption decisions. At the same time, ranchers who lease grazing land from the government will collectively tend to purchase all of the scarce and independently determined number of grazing permits and raise the maximum number of cattle that are allowed by those permits, because it tends only to make economic sense to hold such permits (rather than sell them to another rancher) if one grazes the maximum number of cattle allowed on the relevant parcels of land. A similar upshot emerges even in a more vertically integrated industry such as the poultry industry, where demand is relatively inelastic, and profits are dependent mostly on the cost of inputs such as feed and fuel. See for example the comments of poultry industry expert Ed Fryar in Ryssdal (2015).

As general confirmation of all of this, consider how the Singer/Norcross/Kagan Response would apply to other relevant examples, such as the everyday example of power consumption:

Suppose that there are 250 thousand power consumers in your region, and that each one consumes, on average, 25 units of power per year. Clearly, if only one of those power consumers stopped consuming power, the industry would not respond. Equally clearly, if they all gave up consuming power, vast amounts of power (approximately 6.25 million units per year) would not be produced. But there must also be some number of consumers, far short of 250 thousand, whose renunciation of power would cause the industry to reduce the amount of power produced. The industry may not be able to respond to each individual's behavior, but it must respond to the behavior of fairly large numbers. Suppose that the industry is sensitive to a reduction in demand for power equivalent to 10,000 people giving up power consumption. For each group of 10,000 who give up power, a quarter of a million fewer units of power are produced per year. It appears, then, that if you give up consuming power, you have only a one in ten thousand chance of making any difference to the amount of power produced, unless it is certain that fewer than 10,000 people will ever give up consuming power, in which case you have no chance. Isn't a one in ten thousand chance small enough to render your continued consumption of power inefficacious? Not at all. While the chance that your behavior will have an effect may be small, the effect that is risked is enormous.

As in the original passage from Norcross, the only accurate claim here is the initial observation that the expected effect on production of an individual's consumption is essentially zero, and not of high enough magnitude to equate the expected effect with anything like the average effect. The subsequent argument for the opposite conclusion has to be mistaken, because in the case of power production, as in the case of animal production, we can know enough about the supply chain to know that threshold effects are not sufficiently likely at the margin and are not of sufficient magnitude to drive the expected effect of consumption anywhere close to the average effect on production. For example, although it is easy to imagine that an individual is saving power by turning off the lights in her house, upon reflection it should be clear that the expected effect of such an act on the quantity of power *produced* is essentially zero, because the effect on a large power grid of a single individual turning off power in her home is merely to intangibly reduce the voltage in the neighborhood of that individual in a way that is not measurable from the point of view of power producers, and to increase the reading on her meter in a way that amounts only to a change that is lost in insignificant digits when meter readings are aggregated. As a result, the probability we should assign to an individual having any effect on the quantity of power produced by power plants is vanishingly small in a way that ensures that the expected effect on the quantity produced is closer to zero than it is to the average effect on the quantity produced of all similar actual acts of consumption.

As these examples illustrate, the general forms of argument suggested by the Singer/Norcross/Kagan Response have to be mistaken—for example:

Invalid Argument

Premise 1: You know that some number of (additional) people will perform action A, within some large range of n people, and that the difference made by n (additional) people A-ing would be effect E.

Premise 2: You don't know how many (additional) people will A.

Conclusion: Therefore, the expected E-effect of you A-ing is $(1/n)*E$, bracketing effects from other sources.

This argument is invalid and instances of its conclusion are typically false because we may and usually do have additional information about the mechanisms at play within a collective action situation that requires a very different probability to be assigned to triggering an effect than the simple idea that the probability is $1/n$. As a result, if conclusions such as the one above are to be made plausible, occurrences of ' $1/n$ ' must be replaced with 'the probability of triggering the relevant effects'. However, once that substitution is made, the result no longer supports the Singer/Norcross/Kagan Response, because it does not support anything like, e.g., the claim that the expected effect on animal welfare of one individual becoming a vegetarian is substantially positive on balance. Instead, it supports the initial thought that the expected effect is closer to zero than to the average effect, because given inefficiency, other forms of 'slack' in the supply chain, and the insensitivity of production decisions to the signal generated by a single consumer, it would be unreasonable to assign a probability to triggering a threshold an effect that is sufficiently high to vindicate the Singer/Norcross/Kagan Response.

Analogous remarks apply insofar as the reasoning behind the Singer/Norcross/Kagan is more charitably understood in an attempt to avoid invalidity, by formulating the argument as follows, or by substituting Premise 2* below for Premise 2 above:

Unsound Argument

Premise 1*: You know that the average effect of actions of type A is $(1/n)*E$, given that the aggregate effect of all A-ing is E, and n people are A-ing.

Premise 2*: There is no further evidence available relevant to the expected effect of you A-ing.

Conclusion: Therefore, the expected effect of you A-ing is $(1/n)*E$.

Here the problem is that instances of Premise 2* and instances of the Conclusion are generally false, because, again, there is generally further evidence available that requires a very different probability to be assigned to triggering effects than the argument assumes, for reasons illustrated above in connection with voting and consumer choices. In other words, although it is plausible that perhaps some 'principle of indifference' would tell us to equate the expected marginal effect and the average effect in an imaginary scenario in which we don't know anything except for the average effect of a type of action, in the real world we generally have access to additional evidence that makes it empirically indefensible to equate the expected marginal effect and average effect in such a way, and that makes it similarly

indefensible to assign a probability to making a difference that would be sufficiently high to vindicate the conclusions of the Singer/Norcross/Kagan argument.

The upshot is that when evaluating consequentialist reasons for individual action in collective action situations, the knowledge available about the mechanisms at play in such situations matters greatly, and it is a mistake to think that there is a simple a priori argument that that shows that individuals must always ‘cooperate’ for consequentialist reasons, even when it matters greatly whether we collectively all ‘cooperate’.¹⁴ The crucial point is that the probability of triggering effects that would not otherwise result is a function not only of the number of other actors whose collective action is necessary for such effects, but more importantly also depends on the expected difference between the margin and the tipping points at which those effects can be expected to be generated: and while the probability of triggering such effects is somewhat sensitive to the number of other relevant actors, it is generally much more sensitive to expected differences between the margin and these tipping points. The mistake behind the Singer/Norcross/Kagan Response is to ignore completely the impact of the evidence available about such differences, and to assume instead that we should reason about the probability of making a difference as if there were no reason to think that the signal generated by a single individual will almost certainly be lost in transmission and absorbed by buffers—which is often the opposite of the truth, as investigation of real-world examples such as voting, power production, and livestock production illustrates.¹⁵

In light of all this, we should reject Shelly Kagan’s more general consequentialist analysis of the ethics of collective action. After giving essentially the same argument as Norcross (and Singer) regarding chicken consumption,¹⁶ Kagan writes:

I have discussed the example of purchasing a chicken at considerable length, because I take it to be a fairly representative case of the situation we often find ourselves in with regard to collective action problems. ... But if my discussion of this sort of case is correct, then the consequentialist can handle such cases using the familiar appeal to expected utility. Admittedly, in such cases, I may not be able to know whether or not, if I act, I will be part of a cohort of the relevant size for triggering the bad results. But no matter. I can still know that the expected utility of my act is negative. And that will be enough to allow the consequentialist to condemn my act. ... [As a result,] I believe that [the apparent force of objections such as the inefficacy objection is] misleading. I think that the cases we have been considering do not pose a genuine difficulty for consequentialism. The *collective action problem*—as we might dub it—

¹⁴ Here and elsewhere for ease of exposition I assume a collective action situation in which there is a natural distinction between a ‘cooperative’ collectively-desirable option and a ‘non-cooperative’ collectively-undesirable option, conditional on being chosen by everyone.

¹⁵ Just as given what we know about how elections work, no one should think that there is a simple a priori argument for equating the expected effect of a single individual’s vote with the average effect of everyone’s votes, so too given what we know about how supply chains work no one should think that the Singer/Norcross/Kagan argument is a sound argument for equating the expected effect of a single consumption decision with the average effect of all such decisions.

¹⁶ Kagan (2011, p. 124). See an earlier footnote for references to Singer, Norcross, and others.

can be straightforwardly solved along consequentialist lines. ... there is a small chance that the ['non-cooperative'] act makes a big (morally relevant) difference. And while the chance is only a small one, the difference it makes, if it does make a difference, is sufficiently great to guarantee that the expected utility of the given act is negative. That is the reason the consequentialist can condemn it. ... Collective action cases are all [of this type], though they differ, of course, in their details. ... the consequentialist can handle such cases using the familiar appeal to expected utility.¹⁷

Again, the problem with Kagan's reasoning is that insofar as it generates the intended verdicts on cases, it relies on mistaken assumptions about the probability of an act making a difference, generally mistakenly assuming that $1/n$ is the probability of an act being decisive in real-world collective action problems with n relevant actors and a single threshold effect, or where n acts are required per threshold.¹⁸ However, as we've seen this assumption is not empirically plausible, in part because this assumption implies that there is no information available about the expected difference between the margin and nearest tipping points—which, as we've seen, is not empirically plausible in many real-world collective action cases where our firm considered judgment is that a 'cooperative' act is nonetheless required.

This is a general problem for consequentialist theories, rather than simply a problem for Kagan's particular arguments, because from any consequentialist point of view it is tempting to argue in one way or another for Kagan's conclusions that (a) all collective action cases involve thresholds for morally relevant effects, and (b) when such thresholds are present an appeal to expected utility will always give the verdict that consequentialists desire. In a recent paper, Julia Nefsky has raised important objections to arguments for (a).¹⁹ The argument here is that existing arguments for (b) are unsound, and are ultimately unsalvageable even on the assumption that (a) is correct. In addition to undermining consequentialist theories of the ethics of collective action, this also has more general implications, because as noted above it undermines the expected consequences explanation of such cases that is initially attractive even from a deontological point of view.

In sum, it is impossible for traditional act consequentialism to deliver plausible verdicts on many real-world collective action situations because, by the reasoning above, in large real-world cases individuals often have a probability of making a

¹⁷ Kagan (2011, pp. 129, 111, 120, 140, and 129). Compare Derek Parfit, "Five Mistakes in Moral Mathematics" in Parfit (1984), especially pp. 73–86, where Parfit does not claim that such an appeal to expected utility can explain all of the relevant facts about the ethics of collective action. See also Parfit (unpublished). Parfit's discussion of imperceptible effects cases was inspired by Glover (1975) (see Parfit 1984, fn. 44 p. 511).

¹⁸ For example, see Kagan (2011, p. 124) and the formula Kagan displays on p. 120 regarding the use of $1/n$ as the probability of decisiveness in the base case of a collective action problem with a single threshold.

¹⁹ Nefsky (2011) claims that Kagan's argument for (b) is dubious, but she sets aside the task of explaining in detail exactly how the argument for (b) goes wrong, or if some other argument for (b) might succeed.

difference that is sufficiently small to ensure that ‘non-cooperation’ is the option with the greatest expected value, even when consequentialists themselves agree that ‘cooperation’ is required. The upshot is that appeals to expected consequences cannot offer a plausible account of what individuals are required to do in the kind of collective action situations that are common in modern life, and more generally that the ethics of collective action is more complicated than it initially appears. As the subtle contours of this important area of ethics and social philosophy are further clarified, many of our prior judgments will presumably be vindicated—but a few may also have to be revised.²⁰

Acknowledgements Thanks to an anonymous reviewer, and Chrisoula Andreou, Derek Baker, Alexander Berger, Heather Berginc, Brian Berkey, Tom Blackson, Ted Borer, Cheshire Calhoun, Richard Yetter Chappell, Stew Cohen, Christian Coons, Terence Cuneo, John Devlin, Tyler Doggett, Jamie Dreier, David Faraci, Brian Fiala, Chris Griffin, Liz Harman, Travis Hoffman, Ryan Jenkins, Victor Kumar, Melissa Lane, RJ Leland, Alex Levitov, Jonathan Levy, Hallie Liberto, Eden Lin, Zi Lin, Joel MacClellan, Sarah McGrath, Tristram McPherson, Nathan Meyer, Eliot Michaelson, Julia Nefsky, Alastair Norcross, Howard Nye, Govind Persad, David Plunkett, Joe Rachiele, Rob Reich, Ryan Robinson, Julie Rose, Gideon Rosen, George Rudebusch, Carolina Sartorio, Debra Satz, Dave Schmidt, Dan Shahar, Liam Shields, Daniel Silvermint, Peter Singer, Michael Smith, Patrick Taylor Smith, Dean Spears, John Thrasher, Ian Vandeventer, Chad Van Schoelandt, Alan Wertheimer, Jane Willenbring, Jack Woods, and audiences at the American Philosophical Association Pacific Division, the University of Arizona Center for the Philosophy of Freedom, Bowling Green, and the Colorado State University Animal Ethics Conference for helpful discussions. I am especially indebted to conversations with Harman, McPherson, Michaelson, Plunkett, Reich, and Rosen, and to McPherson’s arguments in his paper “Why I am a vegan”, which greatly influenced my thinking about these issues. For further illuminating discussion of these issues, see Michaelson’s series of posts on veganism and ethics in *The Discerning Brute*, beginning with “Veganism and Futility”, as well as related papers by Harman, Lane, McPherson, Michaelson, and Nefsky.

References

- Brandt, R. (1959). *Ethical theory*. New York: Prentice Hall.
- Brennan, G., & Lomasky, L. (1993). *Democracy and decision*. New York: Cambridge University Press.
- Budolfson, M. (unpublished a). The Inefficacy Objection to Deontology: What it is, Why it is Important, and How to Reply to It, online at www.budolfson.com/papers.
- Budolfson, M. (unpublished b). Collective Action, Climate Change, and the Ethical Significance of Futility, online at www.budolfson.com/papers.
- Dreier, J. (2011). In defense of consequentializing. *Oxford Studies in Normative Ethics*, 1, 97–119.
- Gibbard, A. (1990 [1971]). *Utilitarianism and coordination*. New York: Garland Publishing.
- Glover, J. (1975). It makes no difference whether or not I do it. *Proceedings of the Aristotelian Society, Supplement*, 49, 171–190.
- Kagan, S. (2011). Do I make a difference? *Philosophy & Public Affairs*, 39, 105–141.
- McPherson, T. (2015). Why I am a Vegan. In A. Chignell, T. Cuneo, & M. Halteman (Eds.), *Philosophy comes to dinner*. New York: Routledge.
- Nefsky, J. (2011). Consequentialism and the problem of collective harm: A reply to Kagan. *Philosophy & Public Affairs*, 39, 364–395.
- Norcross, A. (2004). Puppies, pigs, and people. *Philosophical Perspectives*, 18, 229–245.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Parfit, D. (unpublished). What we together do, manuscript dated 29 March 1988.
- Portmore, D. (2011). *Commonsense consequentialism*. Oxford: Oxford University Press.

²⁰ For some judgments that may have to be revised, and an attempt to make further progress on the ethics of collective action, see Rudolfson (unpublished a) and Rudolfson (unpublished b)

- Regan, T. (2004). *The case for animal rights*, updated edition Berkeley: University of California Press.
- Ryssdal, K. (2015). Why chicken wings cost more this time of year. *Marketplace*, 5 January 2015: <http://www.marketplace.org/topics/business/why-chicken-wings-cost-more-time-year>.
- Singer, P. (1980). Utilitarianism and vegetarianism. *Philosophy & Public Affairs*, 9, 325–337.
- Thomson, J. J. (1985). The trolley problem. *The Yale Law Journal*, 94, 1395–1415.