DRAFT – Please do not cite.

Caspar Hare November 2008

Obligation and Regret When There is No Fact of the Matter About What Would Have Happened if You had not Done What You Did[1]

It is natural to distinguish between objective and subjective senses of 'ought'. Roughly: what you ought to do in the objective sense has to do with the merits and de-merits of the options available to you, while what you ought to do in the subjective sense has to do with the merits and de-merits of the options available to you, from your epistemic position. So, for example, when a respectable doctor gives you some pills, it may be (if they are poisonous, though you have no way of knowing that) that subjectively speaking you ought to take them, but objectively speaking you ought to throw them in the bin.

Here are two ways of thinking about the objective ought:

Ought of Omniscient Desire: What you ought^{OD} to do is what an omniscient, rational creature with appropriate interests would want you to do.

Ought of Most Reason: What you ought^{MR} to do is what there is most reason to do.

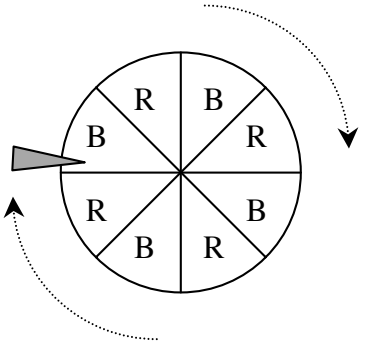
These notions are extensionally different. There are situations in which you ought^{OD} to do one thing but ought^{MR} to do another thing. Or so I will argue in the first part of this paper. In the second part I will look at some useful work to which this distinction can be put.

[1] Many thanks to Bob Stalnaker, Steve Yablo, Agustin Rayo and Ben Bradley for wonderful comments.

1.2 Counterfactual Conditional Under-Specification

The two kinds of ought come apart as a result of *conditional under-specification*. Let me explain what this is.

Imagine a large Wheel of Fortune marked in the manner of a roulette wheel, with red and black spokes.



Imagine that the result of your spinning the wheel from a given starting position is solely determined by the force you apply to the wheel, but that the result is highly sensitive to small variations in that force (the wheel has a sleek aerodynamic profile, its axel is very well-greased, give it a good shove and it will spin for minutes). So the relation between force and result is something like this:

The Force You Apply to the Wheel	**The Result**
.	
.	
15.88345 to 15.88348 N	and you get a RED
15.88349 to 15.88351 N	and you get a BLACK
15.88352 to 15.88356 N	and you get a RED
.	
.	

Finally, imagine that your are offered an opportunity to spin the wheel, but you decline it. Indeed, you never spin a Wheel of Fortune in your life.

What would have happened if you had spun the wheel with 15.88354 N of force? You would have gotten a red. What would have happened if you had spun the wheel with 15.88350 N of force? You would have gotten a black. What would have happened if you had spun the wheel? I say (and I am not being original here) that there is no good, detailed reply to this question. I say, roughly, that because the condition 'if you had spun the wheel' is under-specified, there is no fact of the matter about whether you would have gotten a red, if you had spun the wheel, and no fact of the matter about whether you would have gotten a black, if you had spun the wheel.

How can we put this claim less roughly? That will depend on which semantic theory of counterfactual conditionals we adopt. Consider the counterfactual conditionals:

(1) If you had spun the wheel then you would have gotten a red or a black.

(2) If you had spun the wheel then you would have gotten a red.

(3) If you had spun the wheel then you would have gotten a black.

Adopt David Lewis' theory and (1) is true, but (2) and (3) are false.[2] Adopt Bob Stalnaker's theory and (1) is true, but (2) and (3) have indeterminate truth value.[3]

---

[2] According to Lewis, generally, for propositions A, B, $\ulcorner A \mathrel{\Box\!\!\rightarrow} B \urcorner$ is actually true if a possible world at which $\ulcorner A \wedge B \urcorner$ is true is closer along relevant dimensions of similarity (where the relevant dimensions are fixed by linguistic practice and the context in which the counterfactual conditional is uttered and assessed) to the actual world than any world at which $\ulcorner A \wedge \neg B \urcorner$ is true, and false otherwise. But in this case linguistic practice and the context in which (2) and (3) are uttered and assessed do not fix dimensions of similarity along which a world in which you spin and get a red is closer than any world in which you spin and get a black, or vice versa. See Lewis (1973) sections 1 and 3.4.

[3] According to Stalnaker, generally, for propositions A, B, to assess the truth value of $\ulcorner A \mathrel{\Box\!\!\rightarrow} B \urcorner$ in a certain context, we take that context to fix a closeness relation such that there is a closest world at which A is true. $\ulcorner A \mathrel{\Box\!\!\rightarrow} B \urcorner$ is true if and only if B is true at the closest world at which A is true. But in this case it is indeterminate whether the context fixes a relation such that the closest world in which you spin is one in which you get a red, or a relation such that the closest world in which you spin is one in which you get a black. So it is indeterminate whether you would have gotten a red, if you had spun, and indeterminate whether you would have gotten a black, if you had spun. See Stalnaker (1984) Chapter 7.

For present purposes we need not decide whether Lewis or Stalnaker is right.[4] Let it suffice to note that, according to both Lewis and Stalnaker, (1) is determinately true and neither (2) nor (3) is determinately true. And, according to both Lewis and Stalnaker, for a sufficiently precise condition, like 'If you had spun the wheel with 15.8834 N of force', one of the following counterfactual conditionals

(4) If you had spun the wheel with 15.88354 N of force, then you would have gotten a red.

(5) If you had spun the wheel with 15.88354 N of force, then you would have gotten a black.

is determinately true, the other determinately false. That is all that I will mean when I say that, because the condition 'if you had spun the wheel' is under-specified, there is no fact of the matter about precisely what would have happened if you had spun the wheel.

I take this claim to be uncontroversial.[5] Furthermore, I take it to be uncontroversial that counterfactual conditional under-specification is *resilient under embedding*. Consider the counterfactual conditional:

(6) If you had declined to spin the wheel, then if you had spun the wheel, then you would have gotten a red.

---

[4] The central issue here concerns the principle of Conditional Excluded Middle. How bad is it to deny that if $\ulcorner A \mathrel{\Box\!\!\rightarrow} B \vee C \urcorner$ then either $\ulcorner A \mathrel{\Box\!\!\rightarrow} B \urcorner$ or $\ulcorner A \mathrel{\Box\!\!\rightarrow} C \urcorner$? Stalnaker thinks it very bad. Lewis does not. See Stalnaker (1984) pp. 136-146, and Lewis (1973) section 3.4.

[5] This is not to say that one could not deny it. One could, for example, agree to analyze counterfactuals in a broadly Lewisian way, but insist that features of any given context of assessment determine relevant dimensions of similarity according to which a world in which you spin and get red is closer than any world in which you spin and get black, or vice-versa. We have no way of knowing what the dimensions determined by this context are, but they are there. This would be analogous to epistemicist treatments of vagueness. Or one could say that counterfactual conditionals do not submit to reductive analysis. In addition to all unconditional facts there are primitive conditional facts. One such primitive fact is, e.g, that if you had spun the wheel then you would have applied 15.88347 N to it, and gotten a red. I don't see the appeal behind any such views.

Formally (where 'S' stands for your spinning the wheel and, 'R' for your getting a red, and '$\Box\!\!\rightarrow$' is the counterfactual conditional):

(6) $\neg S \Box\!\!\rightarrow (S \Box\!\!\rightarrow R)$

Supposing that you actually do spin the wheel, and get a red, (6) is not determinately true. Supposing that you actually do spin the wheel and get a red, if you had declined to spin the wheel then there would have been no fact of the matter about precisely what would have happened if you had spun the wheel.

But beware! These uncontroversial claims are easily confused with some independent, highly controversial claims in the realm of metaphysics:

Real Future Contingency: There is no one actual future, only many possible futures.
Nomological Indeterminism: The fundamental laws of nature are indeterministic.

It may be that there is no fact of the matter about what would have happened if you had acted in a certain way, due to the antecedents of the relevant counterfactual conditionals being under-specified, without there being any real future contingency or nomological indeterminism. We did not assume either in the Wheel of Fortune case.

And beware! It is very easy to talk yourself into denying that the antecedents of counterfactual conditionals can be under-specified. After you have declined to spin the wheel, it will sound right to say "I don't know what would have happened if I had spun it." Once you have said that, you may be tempted to infer that there is something to be known that you don't know: either you would have gotten a red, and you do not know it, or you would have gotten a black, and you do not know it. But this is a bad inference.

There is nothing to be known that you don't know. The sense in which you do not know what would have happened is just this: it is not the case that you would have gotten a red and you know it, and it is not the case that you would have gotten a black and you know it. This is the sense in which you do not know the details of a story whose details have never been filled in, the sense in which you do not know the name of Cinderella's birth-mother.

1.3 Deliberative Conditional Under-Specification

So much for under-specification. What does it have to do with what we ought to do? Well, sometimes, when we wonder about whether we ought to do this, that or the other, we think that one thing that matters is what will happen if we do this, that or the other. Ought I to disregard the Queen's orders? I think it matters whether I will or will not be beheaded if I do. I think the truth or falsity of a proposition like

(7) If I disregard the Queen's orders then I will be beheaded

has some bearing the issue at hand.

Now, there has been a disagreement about whether these conditional propositions whose truth or falsity bears on questions about what we ought to do (call them *deliberative conditionals*) are properly understood as future-directed counterfactual conditionals.[6] I will not take sides in this disagreement here. But I will take it that deliberative conditionals, whatever they are, have certain features in common with counterfactual conditionals: First, their antecedents can be under-specified. In the

[6] See Keith DeRose's as-yet-unpublished "The Conditionals of Deliberation".

deliberatively relevant sense of 'if', there is no fact of the matter about precisely what will happen, in the Wheel of Fortune case (in which you do not actually spin the wheel, remember), if you spin the wheel. Second, the under-specification is resilient under embedding. In the deliberatively relevant sense of 'if', supposing that you actually do spin the wheel, then if you don't spin the wheel, then there is no fact of the matter about precisely what will happen if you do spin the wheel.

This raises a question: what ought you to do in cases in which the antecedents of deliberative conditionals are under-specified?

Someone might be tempted to say that there is no need to answer this question, because under-specified options (where an option, O, is under-specified when there is something that matters, M, and there is no fact of the matter about whether O will come about if you take M, due to the antecedent of the deliberative conditional 'If I take O, then M will come about' being under-specified) have no deontic status. When we wonder about what we ought, objectively speaking, to do we should only consider arrays of fully-specified alternatives. In the Wheel of Fortune case, for example, we should not construe the options available to you as [spinning the wheel, declining to spin the wheel], we should construe them as [… applying between 15.88345 and 15.88348 N to the wheel, applying between 15.88349 and 15.88351 N to the wheel, …, declining to spin the wheel]. All of these options are fully-specified, so there's no problem.

But this is not satisfactory. If the proper account of objective oughtness is to have any relevance at all to practical deliberation, the things that have deontic status need to be the sorts of things that it is in our power to do – in the sense that, if we aim to do them, then there is at least a reasonable chance that we will succeed. But it is not in your power

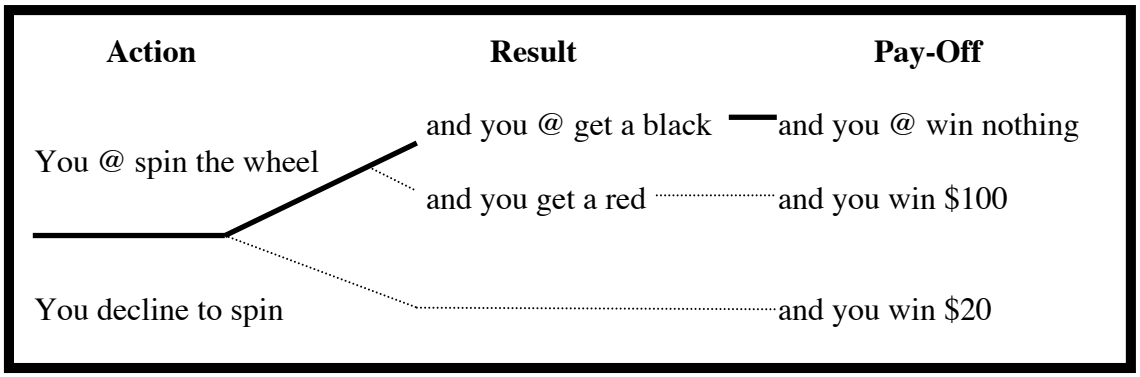
to apply between 15.88345 and 15.88348 N to the wheel. If you aim to do that then you will almost certainly fail. Perhaps it is in your power to do any of the following: [spin the wheel hard, spin the wheel gently, not spin the wheel]. Or even: [spin the wheel really hard, spin the wheel moderately hard, spin the wheel moderately gently, spin the wheel really gently, not spin the wheel]. But we are still a long, long way from the point at which the options available to you cease to be under-specified.

So the question remains: what ought you to do in cases in which the antecedents of deliberative conditionals are under-specified?

1.4 Conditional Under-Specification and the Ought of Omniscient Desire

Let's focus on the ought of omniscient desire. What you ought^{OD} to do is what an omniscient, rational creature with appropriate interests (call her your *Fairy Godmother*) would want you to do. And let's focus on a particular case, in which there are rewards that come with spinning the Wheel of Fortune. If you decline to spin you stand to win $20. If you spin and get a black, you stand to win $100. If you spin and get a red, you stand to win nothing. Prudently, you elect to spin and… you get a red. Shucks!

The Unfortunate Spin



In this case, it is true after the unfortunate spin that

(8) You spun and spinning won you nothing, but if you had declined to spin then you would have won $20.

Your Fairy Godmother will know this, after the spin. So she will presumably wish that you had declined to spin – after all, $20 is better than nothing. So, after the spin, it is true that you ought^{OD} to have declined to spin.

Indeed, supposing that there is no real future contingency, it is true *before* the unfortunate spin that

(9) You will spin and spinning will win you nothing, but if you decline to spin then you will win $20.

Your Fairy Godmother will know this, before the spin. So she will (wistfully, because she knows that her wish is not to be satisfied) wish that you decline to spin – after all, $20 is better than nothing. So, supposing that there is no real future contingency, it is always true that you ought^{OD} to decline to spin.

1.5 Conditional Under-Specification and the Ought of Most Reason

Fair enough. Now let's focus on the ought of most reason. What you ought^{MR} to do is what there is most reason to do. What ought^{MR} you to do in the Unfortunate Spin case?

'Same answer', one might think. 'Consideration (9), which moves your Fairy Godmother to wish that you decline to spin, is a reason to decline to spin – after all, $20

is better than nothing. Furthermore, other things being equal, it is a *decisive* reason to decline to spin, so you ought^{MR} to decline to spin.'

But this is not right. I will try to explain why.

I claim that there is a restriction on the sorts of considerations that are eligible to be reasons for you to do things.

Reasons are not Self-Undermining:

It cannot be that c is a reason for you to ϕ and yet, if you ϕ, c is not true.

Why? Because there is a connection between practical reasons and good practical reasoning. Considerations are eligible to be reasons only if they can play a role in good practical reasoning, only if they are the kind of considerations that can move a good practical reasoner to act. But good practical reasoners are not moved to act by considerations that are not true. That would be bad practical reasoning.

So, for example, suppose that I am deciding between two tedious, mildly unpleasant tasks: mucking out the stables and stripping paint from the roof of the shed. And suppose I know that, if I choose to muck out the stables I will wish (mid-mucking) that I had chosen to strip paint from the roof of the shed, and if I choose to strip paint from the roof of the shed I will wish (mid-stripping) that I had chosen to muck out the stables. And suppose that I actually will choose to muck out the stables. In these circumstances the consideration:

(10) If I strip paint from the roof of the shed, then I will wish that I had chosen to muck out the stables.

may[7] be a reason to choose to muck out the stables. And the consideration:

(11) If I muck out the stables, then I will wish that I had chosen to strip paint off the roof of the shed.

may be a reason to choose to strip paint off the roof of the shed. But the consideration:

(12) I will wish that I had chosen to strip paint from the roof of the shed.

true though it is, cannot be a reason to choose to strip paint from the roof of the shed. Why? Because if I were moved by this consideration to choose to strip paint from the roof of the shed, then I would be reasoning badly. I would be moved by a false consideration.

Now look back at the seemingly decisive reason to decline to spin in the Unfortunate Spin case:

(9) You will spin and spinning will win you nothing, but if you decline to spin then you will win $20.

If you decline to spin then (9) is not true. So (9) is not a reason to decline to spin. Reasons are not self-undermining.

Nor is this a reason to decline to spin:

(13) If you spin then you will win nothing, but if you decline to spin then you will win $20.

[7] I say 'may' because whether or not you think it is a reason will depend on whether you think that considerations concerning our future desires give us reasons to act now. My point here is that it is not *ineligible* to be a reason to act now. Because its truth does not depend on what I do, it is the sort of consideration that can be a reason to act now.

There is space to argue about whether (13) is true in the Unfortunate Spin case, because there is space to argue about whether deliberative conditionals with true antecedents are true just in case their consequents are true.[8] But this is beside the point. Because conditional under-specification is resilient under embedding, if you decline to spin then there is no fact of the matter about whether you will win nothing, if you spin. No matter whether (13) is true, if you decline to spin then (13) is not true. So (13) is not a reason to decline to spin. Reasons are not self-undermining.

This may seem mysterious. How can considerations like (9) and (13) (if true) be reasons to want something to happen, but not reasons to bring it about?[9] Here is a way of dramatizing the idea. We know what a benevolent fairy godmother who believed (9) or (13) to be true would *want* – she would want you to decline to spin, because she believed (9) or (13) to be true. But, supposing that benevolent fairy godmothers could intervene in human affairs, what would a benevolent fairy godmother who believed (9) or (13) to be true *do*? Would she wave her magic wand and make you decline to spin, because she believed (9) or (13) to be true? Not unless she were an imperfect fairy godmother, the sort of fairy godmother who makes practical mistakes. If she waved her magic wand and made you decline to spin because she believed (9) or (13) to be true then she would be acting on a false belief, and to act on a false belief is to make a practical mistake. Her

---

[8] Lewis argued (in Lewis (1973) section 1.7) that we should treat counterfactuals with true antecedents as true just in case their consequents are true. But, as he pointed out, you can easily tweak his semantic apparatus so as to accommodate the denial of this view, by allowing for *weak centering*, allowing for distinct worlds *i*,*k*, such that *i* is as similar to *k* as *k* is to itself. And there is space to argue that, at least in some special contexts, we should go with weak centering.

[9] It is standard (following Parfit, in his as-yet-unpublished but widely read On What Matters) to distinguish between *content-given* and *state-given* reasons for desire. Roughly: state-given reasons for and against desiring that p are considerations that bear on the desirability of desiring that p, while content-given reasons for and against desiring that p are considerations that bear on the desirability of p. Now, there is no mystery in one consideration being a state-given reason to desire that p, but no reason to bring it about that p. But (9) is a content-given reason for desiring that you decline to spin, but no reason to decline to spin. That is prima facie mysterious.

good reasons for wanting you to decline to spin are not reasons for bringing it about that you decline to spin.[10]

What considerations are eligible to be reasons to spin or decline to spin, in this case? – considerations that will be true no matter what you do. For example:

(14) If you decline to spin then you will win $20.

(15) If you decline to spin then there is no fact of the matter about precisely what will happen if you spin, but there is a conditional probability of .5 that you will win $100, and a conditional probability of .5 that you will win $0.

In general, these are considerations concerning the *outcomes* of the actions available to me, where the outcome of the action you do not actually take (call it a_c) is the set of all histories *h*, such that if you had taken a_c, then *h* might have come about, and the outcome of the option you actually take (call it $a_@$) is the set of all histories *h*, such that if you had taken a_c, then if you had taken $a_@$, then *h* might have come about. And on balance such considerations favor your spinning. You ought^{MR} to spin.[11] The ought of omniscient desire and the ought of most reason have come apart!

---

[10] You might object: 'I concede that (9) and (13) cannot move a good practical reasoner to decline to spin, but they are nonetheless reasons to decline to spin – not merely reasons to desire that you decline to spin, but reasons to decline to spin. Cases like the Unfortunate Spin case show that *Reasons are not Self-Undermining* is false. Not all reasons can play a role in good practical reasoning.'

I do not find this way of thinking about reasons very helpful. But if you insist on it, then there is still a question you should find interesting here. Focus on the reasons that can play a role in good practical reasoning. Call them '*P*reasons' if you like. It does not matter so much what you call them. What, on balance, do they favor? What do you have most *P*reason to do? What ought^{MPR} you to do? You should find this question interesting because you acknowledge that good deliberators act on *P*reasons, not on considerations-that-are-not-*P*reasons, and you should be interested in what a good deliberator would do in your circumstances. The answer to the question is surprising: in cases like the Unfortunate Spin case, what you ought^{MPR} to do is not what you ought^{OD} to do.

[11] Jean Paul Vessel, in Vessel (2003), has appealed to an example similar to the Unfortunate Spin example to argue that people concerned with the objective deontic status of actions (he has objective act consequentialists in mind) should reject the view that conditionals with true antecedents are true if and only if their consequents are true. In a context like this the conditional "If you spin the wheel then you will get a black" should be taken to be false, though you will spin the wheel and you will get a black. Vessel's

2. Putting the Distinction to Work

So there are two independent notions: what you ought^{OD} to do and what you ought^{MR}. And in some cases what you ought^{OD} to do is not what you ought^{MR} to do. This raises some questions.

First, there is a venerable tradition in which philosophers analyze the ought-of-most-reason in terms of what fully informed creatures would desire on our behalf. Michael Smith, for example, following Bernard Williams[12], suggests that 'to say that we have normative reason to ϕ in certain circumstances C is to say that, if we were fully rational we would want that we ϕ in C'[13], where being *fully-rational* involves at least having 'all relevant true beliefs' and 'no false beliefs.'[14] But this approach may now appear unpromising. If you had the true belief that spinning the wheel will win you nothing in the Unfortunate Spin case, then you would want that you decline to spin the wheel. But in fact you have most reason to spin the wheel. Must the approach be abandoned?

I think not. We can save the Williams/Smith approach easily enough by saying that being fully rational in their sense involves having all *and only* relevant true beliefs, and placing a restriction on what sorts of true beliefs count as 'relevant' for these purposes.

---

argument relies on a version of the *Principle of Normative Invariance* – which says, roughly, that what you ought to do does not depend on what you will do.

The Principle of Normative Invariance is controversial (Frances Howard Snyder has argued that Vessel's examples should be taken to be counter-examples to the principle – see Howard Snyder (2008) pp. 9-10), is it correct? As I have argued, if the 'ought' that it invokes is the ought-of-omniscient-desire then no. If the 'ought' that it invokes is the ought-of-most reason then maybe. The principle would follow from two claims about reasons: *Reasons are not Self-Undermining* (it cannot be that c is a reason for you to ϕ and yet, if you ϕ, c is not true) and *Reasons are not Self-Supporting* (it cannot be that c is a reason for you to ϕ and yet, if you do not ϕ, c is not true.) Though I think that the argument for the latter claim is less straight-forward than the argument for the former, I am sympathetic to it, so I am sympathetic to Vessel's project, if he is talking about the ought-of-most-reason.

[12] See Williams (1981).

[13] Smith (1994) p. 181.

[14] Smith (1994) p. 156.

True beliefs that would not be true if you were not to do what you actually do (e.g. the true belief that spinning the wheel will win you nothing) do not count.

Second, which notion do we have in mind when we wonder about what we ought, objectively, to do or have done?

I cannot speak for the world here but, for my own part, I find that I sometimes have one notion in mind and sometimes the other. Sometimes I am wondering about whether the action has features that make it the appropriate object of certain conative attitudes (regret, satisfaction, benign indifference). This is typically the salient issue when I wonder, in retrospect – 'Ought I to have taken this risk?' If I discover that it backfired disastrously, then I infer that I ought, in the sense that I am interested in, not to have taken it. The notion I am interested in is Ought^{OD}. At other times I am wondering about reasons to take the action in question. This is typically the salient issue when I wonder, mid-deliberation, 'ought I to take this risk?' I have discovered certain reasons for and against taking the risk. Those reasons, on balance, favor my doing one thing. I am wondering what *all* reasons, discovered and undiscovered, on balance favor my doing. The notion I am interested in is Ought^{MR}.

Now there certainly remains space to argue about which of ought^{OD} and ought^{MR} better deserves the name of 'ought'. But I do not see much future in that argument. I think it will more useful to look at some problems in normative ethics that involve cases relevantly like the Spinning the Wheel case, cases in which the two notions come apart, and see what can be gained from putting the distinction to work.

2.2 Cluelessness

The two notions come apart whenever we care about the consequences of under-specified actions – whenever, due to conditional under-specification, there is no fact of the matter about whether things will turn out well or badly if we do one thing or another.

One sort of person who cares about the consequences of under-specified actions is a *global act-consequentialist*. Roughly: global act consequentialists believe that the objective moral status of an act is determined by the goodness or badness of all its consequences – no matter how remote in space and time.

There is an influential critique of global act consequentialism, due to James Lenman[15], which rests on two claims:

(A) If global act consequentialism is correct then we are, typically, *clueless* about what we ought, objectively speaking, to do. Typically we have good reason to think that there is a big difference in the objective deontic status of the acts available to us, because we have good reason to think that some will have much better consequences than others, but we have no idea what the big difference is.[16]

(B) In situations in which we know ourselves to be clueless about what we ought, objectively speaking, to do, considerations of expected value give us at best 'very weak' reasons to do one thing or another.[17]

[15] In Lenman (2000).
[16] ibid. sections I-III, in particular.
[17] ibid. section IV, in particular.

It would follow that, if global act consequentialism is correct, then we have at best very weak reasons to help old ladies across the street, to refrain from murdering, torturing, raping… and so forth. But that is absurd.

What kind of objective ought is figuring in claims (A) and (B), ought^{OD} or ought^{MR}? Lenman does not say, but I take it that (B) is most plausible when the 'ought' is construed as the ought-of-most-reason. It is hard to get very excited about my reasons for doing something when I know myself to be clueless about whether it is what there is most reason to do.

Is (A) plausible, when the 'ought' is construed as the ought-of-most-reason? To motivate (A) Lenman has us consider a typical prima facie right act: Richard, a 100 BC marauder, spares the life of Angie, a 100 BC village girl. But it turns out that (like many, many 100 BC village girls) Angie was an ancestor of Hitler. If Richard had killed Angie then all the pain and misery that Hitler wrought would never have occurred. It was impossible for Richard to know this, of course. He was clueless about the temporally remote consequences of his act, and hence, if act-consequentialism is correct, clueless about its objective moral status. And even now, with the benefit of two thousand years of hindsight, we are hardly less clueless. We know that sparing Angie led to Hitler, but what would killing Angie have led to? Maybe there would have been another dictator – Malcolm the Truly Appalling – next to whom Hitler looks like little more than a school-yard bully. Maybe there would have been no dictators. We have no idea.

But, when the 'ought' is construed as the ought-of-most-reason, this is not a compelling argument. Consider the counterfactual conditionals:

(16) If Richard had killed Angie then, many centuries later, a dictator named 'Malcolm' would have terrorized the world.

(17) If Richard had killed Angie then, if he had spared Angie, then, many centuries later, a dictator named 'Hitler' would have terrorized the world.

I do not see any reason to think that either is determinately true. The best dynamical models of global weather exhibit extreme sensitivity to initial conditions – small differences in earlier states magnify rapidly in later states. Just as whether I get a red or a black when I spin the Wheel of Fortune depends on whether I spin it with 15.88347 N or 15.88349 N of force, so, according to these models, whether a hurricane strikes Bermuda in November depends on precisely how things are in Massachusetts in May.[18] If our world were exactly the way that these models represent it to be, then a condition like 'If Richard had killed Angie' would be under-specified – how things are many centuries later would depend on *precisely how* Richard killed Angie. (16) and (17) would not be determinately true.

Of course, our world is not exactly the way that these models represent it to be. For one thing, these models represent worlds in which relatively simple laws govern the behavior of a relatively small number of things. In our world, relatively complex laws govern the behavior of a relatively large number of things. For another thing, these models represent worlds with deterministic laws. In our world the laws may not be indeterministic. Well and good, but there is no reason to think that the extra complexity and nomological indeterminism in our world renders counterfactuals like (16) and (17) determinately true.

---

[18] Extreme sensitivity to initial conditions in weather prediction models was first noticed by Edward Lorenz in 1961. See Lorenz (1963). Thanks in large part to his metaphors, it became known as the 'Butterfly Effect'.

Because counterfactuals like (16) and (17) are not determinately true, the relevant comparison, for the purposes of working out what Richard ought^{MR} to have done, is between the set of histories that might have come about if Richard had killed Angie (call this S_{killed}) and the set of histories that, if Richard had killed Angie, might have come about if Richard had spared Angie (call this S_{spared}). In all histories in S_{killed}, Angie is killed, in all histories in S_{spared}, Angie is spared. In some histories in S_{killed}, dictators terrorize twentieth century Europe, and in some they do not. Likewise, in some histories in S_{spared}, dictators terrorize twentieth century Europe, and in some they do not. Prima facie, then, it would appear that Richard ought^{MR} to have spared Angie. Terrible dictator considerations give us reason to doubt this, to think that we are really clueless about what Richard ought^{MR} to have done, only if we have reason to think that, relative to an appropriate measure, there are more terrible dictators in S_{killed} or S_{spared}.

To put this another way: Let 'K' stand for the proposition that Richard kills Angie, 'Sp' stand for the proposition that Richard spares Angie, and 'D' stand for the proposition that a dictator terrorizes twentieth century Europe. And let '$A \Box^{C}\!\!\rightarrow B$' mean the counterfactual conditional probability of B, given A, is C. Terrible dictator considerations would give us reason to think that we are clueless about what Richard ought^{MR} to have done only if we had reason to think that

(17) $K \Box^{M}\!\!\rightarrow D$ and $K \Box\!\!\rightarrow (Sp \Box^{N}\!\!\rightarrow D)$ and $N \neq M$

In English: the counterfactual conditional probability of a dictator terrorizing twentieth century Europe if Richard had killed Angie is not equal to the counterfactual conditional

probability of a dictator terrorizing twentieth century Europe if Richard had spared Angie, if Richard had killed Angie.

But I see no reason to think this, and so no reason to think that we are clueless about what Richard ought^{MR} to have done. He ought^{MR} to have spared Angie.

This is not to say that there are *never* unknown facts of the matter about what we ought^{MR} to do. Suppose you are the captain of a ferry whose boiler blows up mid-passage. As the decks rear up beneath your feet and your passengers slide, screaming, into the frigid waters, you are right to think 'I ought^{MR} to have stayed in port today.' It is rather to say that this is not typically the case. Typically, although you cannot help but be clueless about what will happen as a result of what you do, and hence clueless about whether what you do is an appropriate object of regret or satisfaction, about whether it is what you ought^{OD} to do, you need not be clueless about what there is most reason to do, about what you ought^{MR} to do.[19]

2.3 Parenthood

Another kind of person who cares about the consequences of under-specified actions is a *would-be parent*. Consider a famous example: you are trying to conceive a child. A doctor points out to you that there is a small stretch of time such that, if you conceive a

[19] To be fair to Lenman, I should point out that he does consider the possibility that counterfactuals may be indeterminate in Richard's case. He writes: 'Or perhaps there is no such thing as a determinate way things would have been in the future if he had not acted (or failed to act) as he did. Perhaps such talk of massively complex historical counterfactuals is metaphysical nonsense on stilts and there is nothing here for God to know.' (Lenman (2000) p.252.) But he decides to work on the assumption that this is not so: 'Because all these possibilities make matters even more intractable for consequentialists, let us stick to the most tractable case, in which there are just two possible futures to consider.' I do not immediately see why working with the assumption that the counterfactuals are indeterminate makes things intractable for consequentialists. Certainly if we work with this assumption then we should not take there to be a deep, in principle unknowable fact of the matter about what there is most reason to do, so the problem of cluelessness goes away. I think there are other, serious problems for global act consequentialism, but this is not one.

child during this time then it will almost certainly have health problems. She urges abstinence over this time. You ignore her and, predictably enough, conceive and bear a child, John, who has health problems throughout his life.

A puzzle raised by this example is to explain whether and why you ought not to have ignored the doctor's advice. This is an instance of Derek Parfit's 'non-identity problem', which has received a great deal of attention.[20] A related, but distinct puzzle, which has received less attention,[21] is to make sense of your attitude, years later, towards John's birth.

Years later you have become very attached to John. You are not faking it, when you celebrate his birthday. You are glad that you ignored your doctor's advice and brought John into the world. And so you *should* be. Good parents have this attitude towards their children. And yet you recognize that you ought not to have conceived John. How can this be? How can you simultaneously be glad you did something, and recognize that you ought to be glad that you did it, and recognize that you ought not to have done it?

Here's an explanation: there are two sets of values in play. Compare the state of affairs in which you conceive John to the state of affairs in which you conceive a different, healthy child. Now, being John's parent, you ought to have John-centric values, and ought to prefer the former. Back then, without grounds for attachment to any particular child you might later have, you ought not to have had John-centric values, and ought to have preferred the latter. When you judge that you ought to not to have ignored

---

[20] The canonic presentation of the problem is in Parfit (1984) Chapter 16. For my own views on the problem see my "Voices from Another World".
[21] The clearest presentation of this puzzle that I know of is in an as-yet-unpublished paper by David Velleman: "Love and Non-Existence".

your doctor's advice, you are judging your past actions, not in light of your current values, but in light of the values you ought to have had when you faced those options.

And here's an analogy: I am fighting my enemy. He lunges at me wildly and exposes his left side to my sword. I think 'you ought not to have done that…' and deliver a fatal blow. I am judging his action, not in light of my own preferences (as far as I am concerned, lunging at me wildly was a tip-top thing to do) but in light of his preferences.

But this cannot be the whole story. If it were the whole story then, in order to cast your past action in a negative light, you would need to perform a leap of evaluative imagination, appealing to values that you do not presently endorse. But you do not need to do this. There is a clear sense in which, when viewed in light of your present, John-centric values, and all the facts, you made a mistake by ignoring your doctor: though you ought^{OD} to have ignored your doctor, you ought^{MR} not to have ignored your doctor.

This is because conceiving a baby is a chancy business in the way that spinning a Wheel of Fortune is a chancy business. How it ends is highly dependent on the details of how it started. Without delving too deeply into the goopy details, let it suffice to say that, for any successful conception on your part, it is plausible to think that *that* sperm and *that* egg would not have fused if ten seconds before the conception you had trembled, or shifted your body, or coughed… etc. So the condition 'if you had abided by your doctor's advice and conceived a baby later' is massively under-specified. There is no fact of the matter about precisely what would have happened if you had abided by your doctor's advice and conceived a baby later. And furthermore, since conditional under-specification is resilient under embedding, if you had abided by your doctor's advice, the condition 'if you had ignored your doctor's advice and plunged ahead' would have been

massively under-specified. If you had abided by your doctor's advice, there would have been no fact of the matter about precisely what would have happened if you had ignored your doctor's advice.

So, for the purposes of assessing what you ought^{MR} to have done, we must compare the set of histories that might have come about if you had waited, S_{wait}, to the set of histories such that, if you had waited, then they might have come about if you had not waited, $S_{goahead}$. In most histories in S_{wait} you conceive a healthy child who is not John. In most histories in $S_{goahead}$ you conceive an unhealthy child who is not John. Your John-centric values do not preclude you from favoring S_{wait} over $S_{goahead}$ – after all, John does not exist in most histories in $S_{goahead}$. In light of your John-centric values you ought^{MR} not to have ignored your doctor.

In sum: It is true that an omniscient onlooker who shared your special attachment to John would prospectively wish that you ignore your doctor and retrospectively be glad that you ignored your doctor. But it does not follow that to cast your ignoring your doctor in a negative light you need to set aside your special attachment to John, and adopt a less partial attitude.

2.4 Wrapping Up

Let's review the principal claims I have made here. First, when you think about the objective deontic status of an action there are two questions to ask: 'Is it an appropriate object of regret or of relief?' and 'Is it what there is most reason to do?' Second, it pays to be clear about which question you are asking because sometimes, in situations in which deliberative conditionals are under-specified, the questions have different answers.

The difference arises because, in these situations, reasons to desire that you do something are not reasons to do it, because reasons to act are not self-undermining. Third, this has a bearing on some important problems in normative ethics, because some important problems in normative ethics concern situations in which deliberative conditionals are under-specified. These include situations in which we care about the distant consequences of our actions, and situations in which we care about which children we conceive.

I will close by saying that I find the central idea reassuring. I, like everybody, sometimes take risks. When they backfire I think, regretfully: "In light of what I know now, I wish I had not done that." It is then a short step to the thought: "If only I had known then what I know now, then I would not have done that. It was for want of a measly scrap of information that the bad things happened" And this thought adds a sense of *tragic frustration* to my regret (think of Oedipus unknowingly killing his father, think of Romeo and the friar's letter.) But, reassuringly, the step is often mis-taken. Often my reasons to regret doing what I did were not reasons not to do what I did.

References

Howard-Snyder, Frances (2008): "Damned If You Do, Damned If You Don't!", Philosophia 36, pp. 1-15

Lenman, James (2000): "Consequentialism and Cluelessness", Philosophy and Public Affairs 29, no. 4, pp. 342-370

Lewis, David (1973): Counterfactuals, Oxford: Blackwell

Lorenz, Edward (1963): "Deterministic Nonperiodic Flow", Journal of the Atmospheric Sciences Vol. 20, Issue 2, pp. 130-141

Parfit, Derek (1984): Reasons and Persons, Oxford: Oxford University Press

Smith, Michael (1994): The Moral Problem, Oxford: Blackwell

Stalnaker, Robert (1984): Inquiry, Cambridge, MA: MIT Press

Williams, Bernard (1981): "Internal and External Reasons", in his Moral Luck, Cambridge, England: Cambridge University Press

Vessel, Jean Paul (2003): "Counterfactuals for Consequentialists", Philosophical Studies 112, no. 2, pp. 103-125