How Do Reasons Accrue?

Shyam Nair

0. Introduction

Reasons can interact in a variety of ways to determine what we ought to do. For example, I might face a choice of whether to work on this paper or socialize with friends. And it might be that the only relevant reason to work on this paper is that I have a deadline coming up soon and that the only relevant reason to socialize is that it is relaxing. In this case, whether I ought to work on the paper or ought to stay at home is determined by which of these reasons is stronger.

But there are other cases where we have more than one reason to do a particular act and whether we ought to do that act is not settled by *individually* comparing the strengths of the reasons to do it against the strengths of the reasons to do some alternative. Suppose for example that there is a movie theater and a restaurant across town. And suppose that in order to get to that side of town I must cross a bridge that has a $25 toll. The toll is a reason not to cross the bridge. The movie is a reason to cross the bridge and the restaurant is also a reason to cross the bridge. It may be that if there were just the movie to see, it wouldn't be worth it to pay the toll and if there were just the restaurant, it wouldn't be worth it to pay the toll. But given that there is both the movie and the restaurant, it is worth it to pay the toll.

In this case, the reason to cross the bridge provided by the movie and the reason to cross the bridge provided by the restaurant are individually worse than the reason to not cross the bridge provided by the toll. But together these reasons are better than the reason to not cross the bridge so that you ought to cross the bridge. More generally, cases like this show that we can have two reasons to do an act that are individually worse than a reason to not do that act, but the reasons together are better than the reason to not do that act. So the reasons together——which we can call the *accrual* of those reasons—can have a strength that is an increasing function of the strengths of the individual reasons.

And this way in which can reasons accrue can be seen not just in practical cases concerning what we ought to do but also in epistemic cases concerning what we ought to believe such as the following: I am curious about what color the feathers of a certain bird are. My friend seems to remember reading in a textbook that they are black. I seem to remember seeing in a nature documentary that they are white. I also seem to remember seeing in the travel guide that they are white. It may be that my friend's memory based on the textbook is a better reason to believe that the feathers are black than my memory of the documentary or the travel guide taken individually. But it may be that together these reasons to believe that

the feathers are white are better than the reason to believe that the feathers are black so that I have more reason to believe that the feathers are white. Again, it seems that the accrual of the reasons to believe that the feathers are white can have a strength that is an increasing function of the strengths of the individual reasons.

In this paper, we will look at how reasons determine what we ought to do and believe in cases where the accrual of reasons is relevant. Our focus will not primarily be on questions about the nature of individual reasons and their weight. Instead, we will at the outset rely on our pretheoretical grip on what reasons there are and how weighty they are individually and ask the more formal or structural question of how to determine the strength of their accrual based on these facts. That said, we will be led back at certain points to substantive questions about the nature of reasons and their weight.

In looking at these issues, my goal will not be anything as ambitious as developing a full theory of the accrual of reasons. Rather, my goal will be more modest: I will introducing some of the challenges for providing an adequate model and argue that a promising approach to resolving these challenges involves making use of the familiar distinction in moral philosophy between derivative and non-derivative normative notions.

1. A First Stab

It may be surprising to hear that it is challenging to provide an adequate model of the accrual of reasons. This is because a simple model suggests itself.

Begin with the assumption that the reason-relation is (at least) a three-placed relation holding between a proposition, an action or attitude, and a natural number. The idea is that p is not a reason to φ (e.g., cross the bridge, believe that the feathers are white) simpliciter rather it is always that p is a reason to φ of a certain strength *n*. We then say that one reason is stronger than another just in case the the number associated with the first reason is greater than the number associated with the second.

Finally, we supplement this picture with the following principle of reasons accrual which we may call *Additive Reasons Accrual*:

if
  r_i: p is a reason to φ of strength *m*
and
  r_j: q is a reason to φ of strength *n*
then
  the accrual of r_i and r_j support φ-ing to the strength $m+n$

This principle tells us that the strength of the accrual is an increasing function of the strength of the

individual reasons. And in particular, it says that the strength of the accrual is the sum of the strengths of the individual reasons.[1]

This simple model explains the cases that we began the paper with. Consider the first case. The toll provides a reason to not cross the bridge of some strength *m* (e.g, 25). The movie provides a reason to cross the bridge of some strength *n* (e.g., 20). And the restaurant provides another reason to cross the bridge of some strenght *o* (e.g, 15). It can clearly be that $n < m$ and $o < m$ (e.g, $20 < 25$ and $15 < 25$) so that each reason to cross the bridge is individually worse than the reason to not cross the bridge. Nonetheless it can be that $m < n+o$ (e.g, $25 < 20+15$). So if we accept *Additive Reasons Accrual*, this predicts that the accrual of these two reasons to cross the bridge is stronger than the reason to not cross the bridge as desired. And of course analogous reasoning would allow us to explain the epistemic case.

Though there is much more that could be said about this model and its assumptions, it looks like a promising start toward explaining our opening pair of cases.[2]

2. Its Shortcoming

Unfortunately, there are other cases of accrual that the simple model does not handle well. In particular, there are cases where accrual of a pair of reasons does not have a strength that is an increasing function of the strengths of the individual reasons.

2.1 A Practical Case

We begin with a practical case because these kinds of cases have been most discussed in the literature about our question. Here then is a case slightly adapted from one presented by John Horty (2012: 61):

---

[1] We talk of the accrual "supporting" φ-ing to some strength in order to be neutral about what exactly an accrual is. Though a full theory of accrual would take a stand on this issue, we will be neutral about what an accrual is (e.g, is an accrual itself a reason? if so, is the reason $p \wedge q$?).

[2] On the most straightforward interpretation of this model, it assigns specific natural number to each reason as their strength. This raises at least three issues. First, some basis much be provided for an assignment of a number as the strength of a reasons, The most worked out theory that explains when and why it is appropriate to measure a quantity numerical is measurement theory. So the most promising way to resolve this issue is to use the tools of measurement theory. (See Krantz, Luce, Suppes, and Tversky 1971 for an introduction to measurement theory. I believe that Nozick 1968 was the first to glimpse the connection between reasons and measurement theory.) Second, if each reason is assigned a number and the comparative strength of reasons is determined by the greater than ordering on the natural numbers, it follows that all reasons are comparable. This issue can perhaps be resolved by assigning different kind of numbers for different kinds of reasons or by assigning sets of numbers to reasons. Third, if each reason is assigned a number independently of the other reasons, the strength of reasons will turn out to be an atomic property of a reason rather than a holistic product of the interaction among reasons. This issue can perhaps be resolved by assigning a default strength to each reason rather than a fixed strength. (The relevant notion of *default* can be made precise in default logic (see Horty 2012) and argumentation theory (see Dung 1995). Unfortunately, these theories fit most naturally with qualitative rather than numerical models of weight. And, existing treatments of accrual in these frameworks do not solve all of the problems discussed in this paper, see Delgrande and Schaub 2004, Gómez Lucero, Chesñevar and Simari 2009 and 2013, Modgil and Bench-Capon 2010, Prakken 2005, and Verheij 1995.)

> Suppose I am deliberating about an afternoon run, and that both heat and rain, taken individually, function as reasons to not run; still, the combination of heat and rain together might function as a weaker reason to not run (say, because the heat is less onerous when there is rain).[3]

This is a *prima facie* counterexample to *Additive Reasons Accrual*. The rain and the heat individually provide a reason to not run. We may assume each individually is strong enough to make it that you ought to not run. Nonetheless, the accrual of these reasons supports not running to a lesser degree than the heat and the rain individually support not running. So it may be that in such a case you actually ought to run.

No doubt many readers will already have come up with ways of responding to this (putative) counterexample—and we will spend most of the rest this paper considering these responses—but before we look at them, let's consider an epistemic case. This will give us a firmer counterexample to *Additive Reasons Accrual* and a better appreciation of the full generality of the problem for this account of accrual. It will be important to see the generality of this problem because it will ensure that we develop an adequate *general* strategy for immunizing *Additive Reasons Accrual* from counterexample rather than positing a series of one-off hypotheses to deal with different particular cases.

2.2 An Epistemic Case

Here then is an epistemic case that is a *prima facie* counterexample to *Additive Reasons Accrual*:

> You know that John and Bill are rarely found together—they dislike each other and make it a point to avoid each other. There is a party this week and you are wondering whether John or Bill but not both John and Bill will attend. In this setting finding out John will attend is a reason to believe that John or Bill but not both will attend. Similarly, finding out Bill will attend is also a reason to believe John or Bill but not both will attend.

This is an epistemic case that is a counterexample to *Additive Reasons Accrual*: John's attending is a reason to believe John or Bill but not both will attend. Bill's attending is a reason to believe John or Bill but not both will attend. But their accrual is not a reason to believe John or Bill but not both will attend. After all, 'both John and Bill attend' is inconsistent with 'John or Bill attend but *not both* attend'.

Like the practical cases, there are many ways of responding to this counterexample. But unlike the practical case, we can provide further support for this informal counterexample with the help of some tools from formal epistemology.

2.3 A Formal Interlude

To introduce the formal tools that I have in mind, we need to step away from talking about

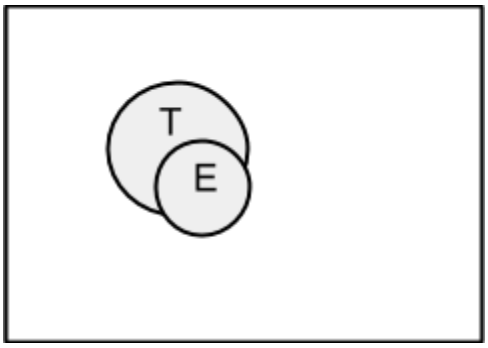
[3] Horty 2012: 2.2.3 and Prakken 2005: §4 present other counterexamples.

reasons for a moment and consider what epistemologists and philosophers of science call *confirmation theory*. Confirmation theory seeks to tell us *whether* a piece of evidence confirms or fails to confirm a given theory and it seeks to tell us *how much* a piece of evidence confirms or fails to confirm a theory.

The standard background framework used in confirmation theory is probabilistic. The idea is that agents have not just all-or-nothing beliefs but also partial beliefs or credences. And the idea is that the credences of a fully rational agent are representable by a probability function.

Confirmation theory uses this framework to ask what is it for a piece of evidence to confirm a given theory. And the epistemologists and philosophers of science who have developed confirmation theory almost unanimously agree that, roughly, a piece of evidence E confirms some theory T for a particular fully rational agent just in case the agent's credence in T goes up when she learns E for certain. More formally, E confirms T for a agent just in case the agent's credence in T, Pr(T), is less than her credence in T conditional on learning E, Pr(T|E).

It may be helpful to visualize this theory as follows: Consider the rectangle below. The points that make up the rectangle can be thought of as possibilities that the agent has some credence in. And a collection of points can be thought of as a proposition, namely the proposition that is true at exactly those possibilities:



Here I have used circles to make vivid two propositions, the proposition that T and the proposition that E. We can think of the agents unconditional credence in T as the portion of the rectangle that T occupies; the area of T divided by the area of the whole rectangle. Similarly for E and other propositions. So just by looking at this diagram we can see that the agent has some non-trivial unconditional credence in T and that the agent has some non-trivial unconditional credence in E and that the agent's credence in T is greater than her credence in E.

We can also use this diagram to see what an agent's conditional credences are. The agents credence in T on E, for example, is represented by the portion of E that T occupies; the area of T that is in E

divided by the area of E. And similarly for other conditional credence. So just by looking at this diagram we can see that the agent's credence in T on E is quite high.

What is helpful about this way of visualizing the model is that allows us to easily "eyeball" our credence and conditional credences. THis makes it easy to determine whether one proposition confirms another. So in the above diagram we can see that E does confirm T because while T does take up some non-trivial portion of the rectangle it takes up a much bigger portion of E.

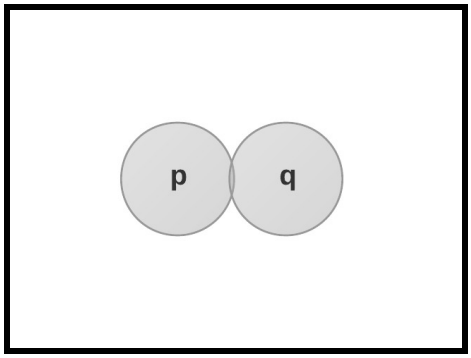
Now as I said, confirmation theory does not just ask *whether* evidence confirms a theory but also asks *how much* it confirms a theory. Unfortunately, the *how-much* question has no uncontested answers. So we will not attempt to settle an answer to this question here.

Now that we have sense of what confirmation theory says and a way of visualizing how the theory works, let us connect this to our discussion of reasons. The connection can be made by noting that that, plausibly, p is a reason to believe q just in case p confirms q.[4] This is especially plausible once we bear in mind that accepting this connection does not force us to accept that reasons can be analyzed in terms of confirmation. Perhaps, the notion of a reason explains the notion of confirmation; perhaps, they are both explained in terms of some third thing; perhaps, they are independent. Whatever the case may be, there are strong connections between them and so it is reasonable to use confirmation theory to shed light on questions about reasons. What's more, even if we do not wish to accept that this connection holds in all cases, we should believe that it holds in many cases and that the class of cases where it does not hold require special explanation.

With all of this in mind, then, we can translate our question about accrual into a question about confirmation: If p confirms r and q confirms r, must p and q together confirm r? Using the formal theory of confirmation, this question comes down to whether the following claim is true: for all p, q, and r, if $Pr(r|p) > Pr(r)$ and $Pr(r|q) > Pr(r)$, then $Pr(r|p \wedge q) > Pr(r)$. As it turns out, the claim is false.

We can illustrate why with a formal version of the counterexample that we gave earlier. Again, we use the points in the rectangle to represent the possibilities the agent has a non-zero credence in. This time I will use two circles to make salient the propositions p and q as can be seen below:

[4] Officially, we only need two (jointly) weaker assumptions for the counterexample: (1) if p extremely strongly confirms q, then p is a reason to believe q and (2) if p is a reason to believe q, then p does not maximally disconfirm q.



Finally, though I haven't made it salient by drawing a circle around it or shading it in a special way, we can also notice another proposition which is the the collection of points that are in the left circle or right circle but not both. Call this proposition r. So r is the non-overlapping portion of the left and right circles; r is the proposition that $(p \vee q) \wedge \neg(p \wedge q)$.

This diagram is a formal version of our counterexample where p is stands for the proposition that John will attend and q stands for the proposition that Bill will attend. Let's check this: First, $Pr(r|p) > Pr(r)$: Pr $(r|p)$ is the portion of the p circle that is not overlapped by the q circle. That is huge because r takes up all of the p space except the tiny part of it that is overlapping q. Pr(r) on the other hand is the portion of the whole space that is either in the p circle or the q circle but not in the overlap of these circles. This is a large portion of the total space but not as large as the portion of the p circle that is not overlapped. Next by analogous reasoning replacing q for p, we can see that $Pr(r|q) > Pr(r)$. So this shows the example satisfies the antecedent of the principle that we are considering.

But it does not satisfy the consequent which, recall, says $Pr(r|p \wedge q) > Pr(r)$. We know the $Pr(r) > 0$ because r takes up some portion of the whole space. But the $Pr(r|p \wedge q) = 0$: r takes up no portion of $p \wedge q$'s space because r is the non-overlapping part of the circles and $p \wedge q$'s space is the overlapping part of the circles.

In fact, though there is controversy about how to measure how much confirmation there is, this kind of counterexample works for all the measures of confirmation that I am aware of. That is, we can show that according to any plausible confirmation measure, the confirmation p provides for r and q provides to r can be arbitrarily close to the maximum (without actually hitting the maximum) while the confirmation provided for r by $p \wedge q$ is the absolute minimum.

The trick is simple. The second feature (that r is confirmed to the absolute minimum by $p \wedge q$) is ensured by the fact that we have chosen r to be logically inconsistent with $p \wedge q$. According to all measures, logically inconsistent propositions maximally disconfirm each other. We ensure the first features (that p and q individually confirm r to arbitrarily close to the maximum without actually hitting it) by

making the overlap of p and q extremely small and making the p-circle and q-circle extremely small in the overall space. It turns out for all the measures that I am aware of, toying with this technique yields the described result.[5]

Thus, in the epistemic case we have intuitive counterexamples to *Additive Reasons Accrual* and can formally verify these counterexamples with the help of some plausible (albeit contestable!) assumptions.

2.4 Responding to the Counterexamples

How should we respond to these counterexamples? There are of course alternative construals of the cases that we may wish to consider (e.g, we might wonder the heat really is a reason to not run rather than the particular discomfort of being in the heat that is a reason to not run). But to adequately motivate these reconstruals and to ensure that we can generalize them to deal with the full range of problematic cases, we do best to consider how to respond to these examples from a more abstract perspective than this. This is especially worth doing given that independently motivated theories like confirmation theory support these counterexamples.

Abstractly then, there are two main ways to respond. One is to claim that these cases are not genuine counterexamples. Another is to accept the counterexamples and to carve out space for a more limited principle of reasons accrual.

The idea of the second is to isolate a class of reasons that do have an accrual that has a strength that is an increasing function of the strengths of the individual reasons. Suppose we give a name to this class, call them *independent* reasons. We then may modify *Additive Reasons Accrual* by adding to the antecedent, the clause that r_i and r_j are independent. Call the revised principle *Additive Independent Reasons Accrual*. By stipulation, *Additive Independent Reasons Accrual* is not subject to the counterexamples presented here.

Of course, this just labels the reasons that avoid our counterexamples 'independent' rather than giving a substantive theory of what these reasons are. In the next sections, we will look at the prospects for giving such a substantive theory. But what makes this labeling a useful piece of bookkeeping is that it allows us to be neutral about whether the first or second strategy is correct. For in the present terminology, the only difference between the first and second strategy is that the first strategy insists that independent reasons are the only reasons that there are whereas the second strategy allows there are other reasons as well but claims only the independent ones have an accrual that has a strength that is an increasing function of the strength of the individual reasons. Thus, no matter which approach we take, it is an interesting question what independent reasons are.

[5] See Fitelson 1999 for discussion of different confirmation measures.

3. A Promising Conjecture about Independence

I know of no systematic way of canvassing all possible substantive theories of independence and in any case, we don't have space to do that here. So instead, what I will do is develop one particularly natural view of independence and argue that it is a promising strategy for solving the problems that we have just encountered.

3.1 Derivative and Non-Derivative Reasons

The key idea of my solution makes use of familiar distinction in moral philosophy between different kinds of normative notions. It is easiest introduce the distinction indirectly by considering the difference between different kinds of good things because this is where this distinction is most often discussed.

Some things are intrinsically or in my favored terminology non-derivatively good and other things are relationally or in my favored terminology derivatively good. For example, plausibly, pleasure is non-derivatively good. Pleasure is good and not because it stands in some relation to something else that is good. Eating ice cream on the other than is plausibly derivatively good. It is good because it stands in some important relation to another thing that is good: for example, eating ice cream causes pleasure. So quite generally, something is derivatively good in virtue of standing in some important relation to something else that is good. And something is non-derivatively good when it is good and not in virtue of standing is some important relationship to something else that is good.

The example of eating ice suggests one important relation that can give rise to derivative good is a causal relation. But we should leave it open what exactly the relations are that give rise to derivative good.[6] For example, a Kantian might think that the sole non-derivative good is the good will. And that pleasure is derivatively good in virtue of standing in certain intentional relations to the good will.

This distinction between derivative and non-derivative goods has been important in moral philosophy. It makes perspicuous the explanatory structured of our theories of goodness. It also allows us to capture subtle difference between goods. For example, it plausible allows us to explain why it is contingent that eating ice cream is good, but necessary that pleasure is good.

And this same important distinction can be applied to reasons as well A derivative reason is a reason in virtue of standing in some interesting relation to some other reason. A non-derivative reason is a reason but not in virtue of standing in such a relation.

My proposal is that independent reasons are non-derivative reasons. And I will argue that this idea

[6] See Korsgaard 1983 and Hurka 2000 who discuss goodness and Väyrynen 2011: 190 and n. 20 for an application to reasons.

is promising because it avoids the counterexample that we saw in the last section.

3.2 Independence in the Epistemic Case

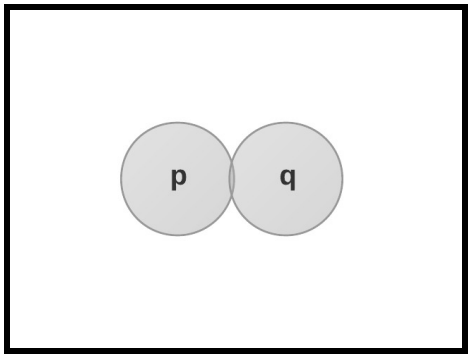
Let's look at the issues involved in epistemic cases first.

To begin, it is helpful to see that the distinction from moral philosophy between derivative and non-derivative reasons is also one that we can easily make among epistemic reasons is. To see this, consider the following case: Someone you know to be reliable tells you that John smokes. This testimony is a reason to believe that John smokes. And it is also a reason to believe that John smokes *or* the number of atoms in the universe is even. Moreover, the testimony is a reason to believe that John smokes or the number of atoms in the universe is even *because* (i) the testimony is a reason to believe that John smokes and (ii) there is a certain relationship between the belief that John smokes and the belief that John smokes or the number of atoms in the universe is even. In particular, the truth of the content of the belief that John smokes ensures the truth of the content of the belief that John smokes or the number of atoms in the universe is even.

In this sense, the testimony provides a derivative reason to believe that John smokes or the number of atoms in the universe is even. The testimony may on the other hand provide a non-derivative reason to believe that John smokes. More generally, this illustrates how derivative epistemic reason can arise based on the logical relationship between contents of beliefs.

Now that we have seen that this distinction among epistemic reasons make sense, we can apply it to our analysis of the counterexample. Recall that the counterexample involved John's attending the party being a reason to think that John or Bill will attend but not both and Bill's attending being a reason to think that John or Bill will attend but not both. These reasons are derivative: John's attending is more fundamentally a reason to believe John will attend and both John and Bill will not attend. The reason to believe that John or Bill will attend but not both is derivative in the same way that the reason in the testimony case is derivative, the logical relationship between the contents of these beliefs—'John or Bill will attend but not both' follows logically from 'John will attend and it is not the case that both John and Bill will attend'.

We can visualize this same point using our diagram:



Recall that the proposition r is the non-overlapping portion of the left and right circles. We saw earlier that p supports r and that q supports r. But *how* does p supports r? We can think of this question as follow: since r is a collection of possibilities, which members of that collection does p increases our credence in? Looking at the diagram, it is not hard to see that p support r by increasing our credence in the possibilities that are in the non-overlapping part of the left circle because the non-overlapping part of the left circle is the place where both p and r hold. How does q support r? q supports r by increasing our credence in the possibilities that are in the on-overlapping portion of the right circle. So p and q each support r in different ways; they increase our credence in different and mutually exclusive "parts" of r. So from this perspective as well, the reasons look derivative. And so this theory of independence immunizes *Additive Independent Reasons Accrual* from the epistemic counterexample.

3.3 Independence in the Practical Case

I also believe that this account of independence can help in the practical case. As we said one standard places where this distinction arises in practical cases is when there are causal relations involve. So to illustrate, suppose that I have promised Mark that I will meet him for lunch downtown and suppose that the only way to get downtown is to take the train. In this case, there is a reason to meet Mark for lunch and there is a reason to take the train. The reason to take the train is derivative of the reason to meet Mark for lunch in virtue of the fact that taking the train is a necessary means to meeting Mark for lunch.

Before showing how this distinction avoids the counterexample that I gave above (which I will do in a moment), it is worth seeing that it allows us to predict fresh counterexamples to *Additive Reasons Accrual*. The fact that we can predict these counterexamples is corroborating evidence that the theory of independence that we are considering is on the right track.

To see the counterexample we need only consider a variant of our original case. As before, the toll is $25 dollars and as before, there is a restaurant and a movie theater that I can access only by paying this toll. But in this case, let's suppose that the movie only has one showing and the restaurant only has one

seating and they are at the same time so that I cannot attend both. Still, the movie is a reason to cross the bridge and the restaurant is a reason to cross the bridge. But the accrual of these reasons is not any stronger than these reasons individually.

This is telling counterexample to *Additive Reasons Accrual* because it is clear why the accrual of the two reasons to pay the toll is not stronger than the individual reasons: The reasons to cross the bridge are *derivative* reasons because they are reasons to take the means to ends that I already have reasons to pursue, seeing the movie and eating at the restaurant respectively. And in this case, these ends are incompatible with one another.

Contrast this with a case where I will enjoy the movie and I have promised to attend the movie. In this case, the accrual of these two reasons to see the movie does have a strength that is greater than the strengths of these two individual reasons. And the reasons to see the movie also are intuitively not derivative. This corroborates our conjecture that independent reasons are non-derivative reasons.

Finally, return to the running case. Here we said that the heat is a reason to not run and the rain is a reason to not run. *Additive Independent Reasons Accrual* applies to this case according to the account of independence that we are considering only if these reasons are non-derivative.

Now it will be controversial what the non-derivative reasons in this case are (a point that we will return to in a moment). But on any plausible theory it is not the mere fact that it is hot that non-derivatively supports refraining from running or the mere fact that it is raining that non-derivatively supports refraining run. Perhaps, the fact that it is raining non-derivatively supports not being out in the rain or something along these line. But whatever the precise reason may be, the rain and the heat do not non-derivatively support refraining from running and so this case is not a counterexample to *Additive Independent Reasons Accrual.*

Thus, if we interpet independent reason as non-derivative reasons, we avoid the counterexamples given above. So this account of independent reasons is a promising strategy for developing a model of the accrual of reasons.

4. The Limits of the Strategy

Unfortunately, there are still challenges that remain for this model. I will present two.

4.1 Derivative/Non-Derivative and Confirmation

So far, I have introduced the derivative/non-derivative distinction and then relied on our pretheoretical judgments to determine whether a given reason is derivative or not. The first problem concerns whether this pretheoretical appeal to the derivative/non-derivative distinction can be safely relied on.

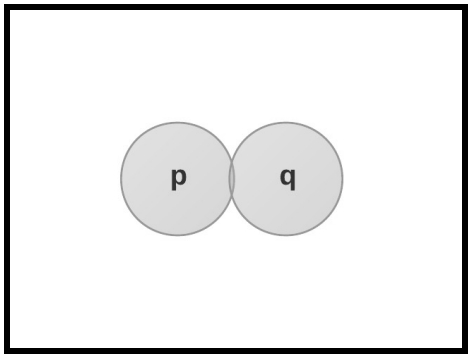
We have already seen when discussing the running case that it is not clear what the non-derivative reasons are. In fact, different theories of practical reason make different claims about this case. So the fact that it is controversial what the non-derivative reasons are is one source of worry about only relying on our pretheoretical judgments.

But another, perhaps sharper, reason to worry is that this distinction fits uncomfortably with confirmation theory. In confirmation theory there is no distinction between derivative and non-derivative confirmation. The theory tells us about confirmation simpliciter.

Of course, I've already argued that there is some distinction between derivative and non-derivative reasons in the epistemic case. So everyone must allow for it. But what looking at confirmation theory points out is that we have worked out theories that do not make use of this distinction.

So we must do the hard work of showing how these theories can be made to fit with this distinction and then check whether the predictions that these theories make once they incorporate this distinction get us the results that we want. Or at least, this is what would be required if we wish to be on firmer footing in appealing to the derivative/non-derivative distinction to solve our problem

Here then is a suggestion on how to fit the the derivative/non-derivative distinction within confirmation theory. Look back at our diagram:



As I pointed out earlier, p confirms the non-overlapping portion of both circles by confirming the non-overlapping portion of the left circle This suggests p non-derivatively supports $p \wedge r$. Analogous considerations suggest that q non-derivatively supports $q \wedge r$. My suggestion would be to try to generalize and formalize these observations about this particular case. Ideally, we would like to show that we can recover full-fledged confirmation measures from a non-derivative measure of confirmation and a theory of the derivative confirmation. This is an open problem that we must solve before we have a full theory of the accrual of reasons.[7]

[7] Though I do not have the space to present the proposal here, in work in progress I show that one can define a certain notion of one reason not being derivative of another reason in probabilistic terms (in particular in terms of whether certain conditional credences

4.2 Finding Your Accrual Partners

The second problem is that this account may purchase its immunity to counterexample at the cost of losing its explanatory power.

To see this, return to our first case. There we had two reasons to cross the bridge that were individually worse than the reason against crossing the bridge. But as we pointed out the reasons to cross the bridge are derivative.

These reasons then do not additively accrue according to *Additive Independent Reasons Accrual* which recall says this:

if
    r_i: p is a reason to φ of strength m
and
    r_j: q is a reason to φ of strength n
and
    r_i and r_j are independent
then
    the accrual of r_i and r_j support φ-ing to the strength $m+n$

In order to use *Additive Independent Reasons Accrual* to explain this case, we would need to find some non-derivative reasons and these reasons would need to be reason to do *the very same act*. The trouble is that it is not at all promising that there are such non-derivative reasons to do the very same act around.

As we already pointed out, it is hard to be sure exactly what the non-derivative reasons in a given case are. But roughly, it seems that the features of the restaurant non-derivatively support something like eating at that restaurant and the features of the movie non-derivatively support something like watching the movie. So there is no single act that the restaurant and the movie both non-derivatively support.[8]

We must then find a new way of determining which reasons additively accrue. In particular, we must somehow allow reasons to do different acts to accrue.

And there is some reason to be optimistic that we can determine which reasons additively accrue. To begin, our counterexample suggests a minimal constraint. We saw that the reason provided by the

---

the agent has are probabilistically independent). Given this definition, one can prove that reasons of this sort always have accruals whose strength is an increasing function of the strengths of the individual reasons and prove that the strength is in fact the sum of the strength of the individual reason (assuming that the strength of a reason is measured by what is known as the log-likelihood measure of confirmation).

[8] This is too quick. For example, we could say that there is a non-derivative reason to do pleasurable things and both the reasons are derivative of this reason. At the limit, this strategy settles the answer to all questions involving reasons and their weight. The idea would be that there is just one non-derivative reason which is to have pleasure and that the weight of a derivative reason is determined by how much pleasure the act provides.

I do not wish to rely on a simplistic theory like this for two reasons. First, it is controversial. Second, one of the most attractive features of explaining what we ought to do in terms of reasons is that it allows for pluralism about which things are normatively significant. See the introduction to this volume for further discussion.

restaurant and the movie don't additively accrue when you can't both go to the restaurant and the movie. So a natural first pass account says that non-derivative reasons to do acts that can be done together additively accrue.

Now this first pass may not be enough. This is because there are cases where we can do two acts together but doing one of the acts makes it harder to do the other. In this kind of case, it is not obvious that reasons additively accrue. In order to solve this problem, more work will be needed. But a natural next step to take to solve this problem is to enrich our model so that i can represent how much harder doing one acts makes doing another and then weight the strength of the accrual according to how much harder doing the first act makes doing the second. One way to do this is to introduce probabilities and take the sum of the strength of the reasons weighted by the probability of doing the second act given that you do the first one.

Thus, though the account of independence on offer faces some challenges, these challenges appear to be surmountable. I conclude therefore that a model along the lines that I have been developing is a promising one for understanding accrual even though there are a number of details of this model that still must be worked out. [9]

5. What's Wrong with Bruteness?

Let me close this paper by consider a more pessimistic perspective on the progress that we have made so far. According to this perspective, the fact that there are issues to still be worked out and the fact that we do not yet know for sure whether they can be worked out should make us skeptical that a productive theory of independence can be had. From this perspective, it may be tempting to treat independence as brute. The idea would be that there is nothing helpful to say about independence. And perhaps this is not so implausible because we can judge in particular cases that certain reasons add up and this shows that we can tell which reasons are independent without a theory. According to this view, *Additive Independent Reasons Accrual* and the simple model are true. But there is nothing helpful to say about independence.

While this is a rather blunt account of independence, it is one that many people have been drawn to.[10] So it is worthwhile to close this paper by discussing what is unsatisfying about treating independence as brute.

[9] Recent work by Bader in this volume and Wassell ms addresses issues similar to the ones discussed here though each sets up the problem of reasons accrual in his own distinctive framework.

[10] Confirmation theory, the models mentioned in n. 2, Pollock 1995: 101-02, Ross 1930, and Schroeder 2007: ch. 7 all treat (aspects of) accrual as brute. Scanlon 2014: ch. 5 also appears to be open to this (but see Scanlon 2014: 107-108, Schroeder 2014 and the introduction to this volume for some complications).

5.1 Ross

One way to bring out what is unsatisfying about this strategy is to consider a classic objection to Rossian moral theory (Ross 1930). W. D. Ross thought that we had a number of *prima facie* duties. In standard cases, more than one of these duties apply and support incompatible actions. What we ought to do is one of these actions and which action we ought to do, according to Ross, is determined by the *prima facie* duties that apply. But Ross thought there could be no helpful theory of how these *prima facie* duties come together to determine what we ought to do.

A classic objection to Ross's view has been that it is unacceptably brute to claim that *prima facie* duties determine what we ought to do while insisting that there can be no helpful theory of this. The worry is that to treat how prima facie duties determine what we ought to do as brute is as unsatisfying as simply treating facts about what we ought to do as brute.

Obviously, this same worry in slightly different terminology arise for treating reasons accrual as brute.

5.2 Generalizations

Another problem is that there are a number of generalizations about reasons accrual.[11] This does not sit easily with the brute approach because it must take the generalizations as brute rather than predicting or explaining them. To illustrate, I will present one of the several generalizations that theorists have uncovered.

To state this generalization, we need some terminology. Philosophers typically distinguish between *undercutting defeat* and *rebutting defeat*. We can introduce this distinction with a pair of examples. In the first case, you know John is generally reliable and he tells you that water is H_20. You know that Sam is even more reliable than John and he tells you that water is not H_20. Plausibly, the reason provided by Sam's testimony is stronger than the reason provided by John's testimony so that you shouldn't believe that water is H_20. In the terminology, this fact is indicated by saying that Sam's testimony is a *rebutting* defeater of John's. In general, when we have two reasons a necessary condition of one being a rebutting defeater of another is that the reasons support incompatible conclusions.

Contrast this case with the following one. As before, you know that John is reliable and he tells you that water is H_20. But in this case, you are told by Sally who you know is generally reliable that while John is generally reliable, he is not reliable about issues related to chemistry. In this case too, it may be that Sally's testimony makes it so that you should not believe that water is H_20. But unlike the case involving Sam, Sally's testimony is not a reason to believe water is not H_20. Rather it just "attacks the connection"

[11] See Prakken 2005: §3 for a list.

between John's testimony and the belief that water is H_20. In the terminology, this fact is indicated by saying that Sally's testimony is an *undercutting defeater* of John's testimony. In general, when we have two reasons and one is an undercutting defeater of another, it need not be that the reasons support incompatible conclusions. Rather the undercutting defeater attacks the connection.

In this terminology, our original case illustrates that a reason can be a rebutting defeater of two reasons, but not a rebutting defeater of their accrual. But a striking fact is that there do not appear to be cases where a reason is an *undercutting* defeater of two reasons, but is not an undercutting defeater of their accrual.

It is of course difficult to prove generalizations like this. But it is at least hard to find cases where this generalization is false. And that's an interesting fact about the accrual of reasons. Perhaps, it is a brute fact. But it would be better if we could predict this general fact on the basis of a systematic account of the accrual of reasons.

5.3 A Flat-Footed Complaint

Finally and most flat-footedly, let me register my incredulity at the thought that the strength of the reason to pay the toll in the case where there is a restaurant and a movie to see on the other side somehow floats totally free of the reasons provided by the restaurant and the reason provided by the movie or that the connection between these reasons is one that we can say nothing helpful about. When we look at the case, this seems obviously false. And similar sentiments apply to other cases that we have discussed.

Of course, none of the considerations are decisive. But they do suggest that the brute approach is not the most palatable option. Providing a systematic theory would be better.

6. Conclusion

I began this paper by presenting cases involving the accrual of reasons and by asking how best to understand these cases. In trying to answer this question, I started out by relying on our pretheoretical grip on what reasons there are and how weighty they are individually. I chose to do this in order to highlight directly the challenges for understanding the accrual of reasons. The main challenges was to develop a productive account of independence and my argument has been that a promising way of developing such an account relies on the familiar distinction in moral philosophy between derivative and non-derivative reasons. But we also saw that to fully answer this challenge and develop a complete theory, we will need to delve further into theoretical questions about the nature of reasons and their weights.

This illustrates how in order to fully understand accrual, we must tackle a tangle of interesting structural and substantive questions about reasons. We have made some progress toward answering these questions. But there is much work left for epistemologists, ethicists, logicians, and others interested in the

weight of reasons to do.[12]

[12] For discussion of these issues, I thank the participants of the USC dissertation seminar, an anonymous referee, Alex Dietz, Kenny Easwaran, Ben Lennertz, Matt Lutz, Barry Maguire, Kenny Pearce, Doug Portmore, Mark Schroeder, Damien Wassell, Jonathan Wright, and especially Errol Lord.

References

Bader, Ralf. this volume. "Conditions, Modifiers and Holism"

Dancy, Jonathan. 2004. *Ethics without Principles*. Oxford: Oxford University Press.

Delgrande, James and Torsten Schaub. 2004. "Reasoning with Sets of Defaults in Default Logic" in *Computational Intelligence* 20: 56-88.

Dung, Phan Minh. 1995. "On the Acceptability of Arguments and Its Fundamental Role in Nonmonotonic Reasoning, Logic Programming, and *n*-Person Games" in *Artificial Intelligence* 77: 321-357.

Fitelson, Branden. 1999. "The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity" in *Philosophy of Science* 66: S362-S378.

Gómez Lucero, Mauro, Carlos Chesñevar, and Guillermo Simari. 2009. "Modelling Argument Accrual in Possibilistic Defeasible Logic Programming" in *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*: 131–143.

—. 2013. "Modelling argument accrual with possibilistic uncertainty in a logic programming setting" in *Information Sciences* 228: 1 –25..

Hurka, Thomas. 2000. *Virtue, Vice, and Value*. Oxford: Oxford University Press.

Korsgaard, Christine. 1983. "Two Distinctions in Goodness" in *Philosophical Review* 92: 169-195.

Krantz, David, Duncan Luce, Patrick Suppes, and Amos Tversky. 1971. *Foundations of Measurement*, vol 1. New York: New York Academic Press.

Modgil, Sanjay and Trevor Bench-Capon. 2010. "Integrating Dialectical and Accrual Modes of Argumentation" in *Proceedings of the 2010 Conference on Computational Models of Argument*: 335–346.

Nozick, Robert. 1968. "Moral Complications and Moral Structures" in *Natural Law Forum* 13: 1-50.

Pollock, John. 1995. *Cognitive Carpentry*. Cambridge: MIT Press.

Prakken, Henry. 2005. "A Study of Accrual of Arguments" in *Proceedings of Tenth International Conference on Artificial Intelligence and Law*: 85-94.

Ross, W.D. 1930. *The Right and the Good*. Oxford: Oxford University Press.

Scanlon, T.M. 2014. *Being Realistic about Reasons*. Oxford: Oxford University Press.

Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford: Oxford University Press.

—. 2014. "*Being Realistic about Reasons*, by T.M. Scanlon" in *Austrialasian Jounral of Philosophy* .

Väyrynen, Pekka. 2011. "A Wrong Turn to Reasons?" in Michael Brady ed. *New Waves in Metaethics*. New York: Palgrave. 185-207.

Verheij, Bart (1995). "Accrual of Arguments in Defeasible Argumentation" in *Proceedings of the Second Dutch/German Workshop on Nonmonotonic Reasoning*: 217–224.

Wassell, Damian. ms. "When Are Accruals of Reasons Stronger than Their Elements?"