The agent-relative/agent-neutral distinction: my two sense(s)[1]

Jessica Lerm

Abstract

The agent-relative/agent-neutral distinction is very well established and widely employed in the metaethical literature. However, I argue that there are actually two different senses of the distinction at large: the hetero-/homogeneous sense and the dependence/independence sense. The traditional, unqualified distinction ought, therefore, to be amended, with each use of the distinction being stipulated as used in either the hetero-/homogeneous sense or the dependence/independence sense. Careful analysis of various metaethics supports that there are these two senses – analysis, in particular, of a neo-Kantian metaethic, according to which reasons are agent-relative in the dependence sense but agent-neutral in the homogeneous sense, and – perhaps surprisingly – of Utilitarianism, according to which reasons are agent-neutral in the independence sense but agent-relative in the heterogeneous sense.

Introduction

Minimally, a reason is a consideration in favour of some action.[2] This is a normative definition, such that, if an agent has a reason to Φ and he/she fails to Φ, he/she has done something *wrong* – although the definition leaves open the question of whether this is moral wrongness or merely prudential, etc., wrongness. This very basic definition of a reason also leaves open the question of a reason's defeasibility; that is, the definition allows that (1) if an agent has a reason to Φ as well as a reason to not-Φ, his/her reason to Φ may be entirely defeated by the greater reason to not-Φ, or (2) his/her reason to Φ may simply be trumped but not defeated by the greater reason to not-Φ. It also leaves open the question of (motivational) internalism/externalism; that is, the definition allows that (3) if an agent has a reason to Φ, his/her having a reason entails that he/she is at least minimally (defeasibly) motivated to Φ, or that (4) the agent may be subject to a reason to Φ and yet have no motivation to Φ at all. Anything more substantial than this is difficult to say without committing to the nature of reasons as either agent-relative or agent-neutral, and, the direction in which the metaethicist decides – agent-relative, or agent-neutral – will bring with it a host of further implications about reasons.

[2] This definition precludes theoretical reasons. This paper deals with the agent-relative/agent-neutral distinction as applied – as it most commonly is – to metaethics, and thereby to practical, and not theoretical, reasons.

The agent-relative/agent-neutral distinction is very well established in the literature. In 1970, Thomas Nagel drew a distinction between 'subjective' and 'objective' reasons (Nagel 1970: p. 90). In 1984, Derek Parfit drew a distinction between 'agent-relative' and 'agent-neutral' ethical theories, which he identified with Nagel's original subjective/objective distinction, writing: 'Nagel's subjective reasons are reasons only for the agent. I call these agent-relative [...]' (Parfit 1984: p. 143, as cited by Ridge 2005: s. 1; italics removed). Two years later, Nagel himself (1986) adopted the newer Parfitian terminology of 'agent-relative'/'agent-neutral'. Since then, the agent- relative/agent-neutral distinction has pervaded the metaethical literature, establishing itself as a standard tool in the limited arsenal we have available to help us settle metaethical debates about the nature of reasons. However, as I hope to demonstrate in this paper, I think the agent-relative/agent-neutral distinction has outlived its usefulness.

Here follows a small sample of some definitions and descriptions of the agent-relative/agent- neutral distinction from the literature. The sample is taken from metaethical main-players Nagel, Parfit and Korsgaard – however, since the distinction is so pervasively employed in the meta- ethical literature, my aim in this paper is to critique the *distinction* itself, rather than any particular philosopher(s), so I take what follows merely as a sample of a more general trend in the literature.[3][4][5][6][7]

	Agent-relative	Agent-neutral
A[3]	The general formulation of which contains an essential reference to the person who has it	Can be formulated such that it does not include essential reference to the person who has it
B[4]	Gives different aims to different agents	Gives common (moral) aims to all agents
C[5]	Exists only in relation to individuals	'There for everyone'
D[6]	A reason only for a particular agent to promote something but not necessarily for others	A reason for anyone to promote the thing
E[7]	Personal property	Common property

Definition A is taken from Nagel (1986), B from Parfit (1984), and C–E from Korsgaard (1993). As mentioned above, Nagel and Parfit profess to be marking out the same distinction. However, it's far from obvious that definitions A and B draw the same distinction. Definition A seems, rather, to be saying something more like C than like B; and definition B seems, rather, to be saying something more like D and E[8] than like A. However, as mentioned, my aim in this paper is not to critique either Nagel or Parfit in particular, but to critique instead the agent-relative/agent-neutral distinction itself, given how pervasive the terminology has become following its

---

[3] Nagel (1986: pp. 152–153).
[4] Parfit (1984: p. 27), as cited by Ridge (2005: s.1).
[5] Korsgaard (1993: s. Introduction).
[6] Korsgaard (1993: s. I).
[7] Korsgaard (1993: s. I).
[8] That definition B is saying something (more) like E may require explanation. The nature of 'personal property' seems to preclude the possibility of you and I owning the same thing. For instance, if it's my car, it can't be your car. So you and I must have different cars. Likewise, if it's my reason (E), it can't be your reason. So you and I must have different reasons/aims (B).

2

popularisation and (I will argue) confusion. As just one example, consider how all three definitions C, D and E issue from the same one philosopher – yet these three definitions are not obviously equivalent. Compare, for instance, C's and E's definitions of the agent-relative. According to C, agent-relative reasons exist only in relation to individuals; and, according to E, agent-relative reasons are private property, possessed only by the agent. However, E's notion of private property doesn't appear to be consistent with C's claim about existence. The notion of private property seems to presuppose that the possession has independent existence – *contra* C – such that it could have belonged to you, but, as it so happens, it actually belongs to me. Take, for instance, my car: prior to its coming to be mine, it existed independently in the car lot; you could have bought it, but you didn't; instead, I bought it, and now it's my private property (E). However, this certainly doesn't mean that my car exists only in relation to me (C). Thus, even though they go by the same names of 'agent- relative'/'agent-neutral', there appear to be various different definitions of this distinction at large in the literature.

I wish to argue that there are (at least[9]) two distinct senses of 'agent-relative'/'agent-neutral' portrayed in this sample, with A and C drawing one kind of distinction, and B, D and E another. I go on to defend the claim that there are these two senses of the distinction against the objection that there is, in fact, just one distinction and that the apparent differences in definitions are only superficial – that is, the 'reduction objection'. I mount my defence against the reduction objection by appealing to a variety of 'combination cases'. A combination case is a seriously plausible metaethic according to which a reason is agent-relative in one sense but agent-neutral in another. Given that the traditional, unqualified distinction obscures these combination cases, which are seriously plausible metaethics, my contention is that either, modestly, we ought to qualify each future use of the terminology by stipulating the sense in which it is being employed or, more aggressively, the traditional terminology ought to be abandoned entirely, in favour of a vocabulary less confusing. These two senses of the agent-relative/agent-neutral distinction are discussed below.

The two senses
The hetero-/homogeneous sense
Call the first sense of the traditional agent-relative/agent-neutral distinction the *hetero-/ homogeneous* sense of the distinction. Take definition B, above, as the standard definition of the agent-relative/agent-neutral distinction on this hetero-/homogeneous sense. According to this sense of the distinction, reasons are agent-relative when they differ from person to person. For instance, if I have a reason to go on a diet, but you don't have a reason to go on a diet, then that reason, which holds differentially across you and me, must be an agent-relative reason in the heterogeneous sense. On the other

---

[9] There may well be further variant interpretations of the traditional terminology than the two I go on to describe. Prof. Thaddeus Metz has suggested, for instance, that 'agent-relative'/'agent-neutral' might also be used to indicate the weakness/strength or defeasibility/primacy of a reason, and I hint at a further alternative interpretation in fn.24. Unfortunately however, I am unable to explore fully this possibility here. However, of course, the more distinct senses of the traditional distinction there are, the stronger my argument against it.

hand, reasons are agent-neutral in this hetero-/homogeneous sense of the distinction when they do not differ from person to person but hold uniformly across persons. For instance, you might think that everybody has a reason to floss daily. Of course, hardly anybody actually *does* floss daily, but you might think that, insofar as daily flossing is crucial to dental hygiene, which is crucial to one's health, everybody has a reason to floss daily, and this reason holds uniformly across persons. Moral reasons are also a popular candidate here:[10] you might think that everybody, no matter who you are – the richest of the richest or the poorest of the poor – has a reason to, say, help others when it's at very little cost to themselves. If there's a child drowning in the pool, and you can swim, it doesn't matter if you're the president of the universe or just the milkman: you have a reason to jump into that pool and save the drowning child. Reasons such as this, that hold uniformly across persons, are homogeneously agent-neutral; they're agent-neutral in the sense that they apply neutrally across all persons, not 'favouring' some over others.

A primary example in ethics of the agent-relative/agent-neutral distinction in this hetero-/ homogeneous sense would be the contrast between Egoism and Deontology. Egoism is straightforwardly heterogeneously agent-relative in that it gives different egoistic reasons to difference egos, or agents. Deontology, on the other hand, is homogeneously agent-neutral in that its reasons are categorical, applying across all agents, in all circumstances, absolutely.

The dependence/independence sense
Call the second sense of the distinction the *dependence/independence* sense. According to this sense, reasons are agent-relative when they are dependent on the agent in some sense, when the relationship of reason to reason-holder is essential to the reason's being. On the other side of the divide, reasons are agent-neutral in this sense when the relationship of reason to reason-holder is merely contingent, when the reason could exist independently of the subject to whom it applies. Whereas homogeneously agent-neutral reasons (above) apply uniformly to everybody and are neutral amongst agents, independently agent-neutral reasons are neutral *of* agents. Unlike the former hetero-/homogeneous sense, this dependence/independence sense is a *metaphysical* take on the distinction.

A perceptual analogy will help to see the difference between the agent-relative and the A perceptual analogy will help to see the difference between the agent-relative and the agent-neutral in this dependence/independence sense. Consider an object – say, this pen. This pen has mind-independent existence. If I close my eyes, or turn around, or leave the room, the pen will still be there. If I drop dead, the pen will go on existing, for a time. However, consider now your perceptual impression of the pen itself. This perceptual impression, unlike the pen itself, is mind-*dependent*: if I dropped dead and stopped existing my perceptual impression of the pen would also stop existing, even though the pen itself would go on. On the dependence/

---

[10] Indeed, as definition B suggests, Parfit (1984) apparently identifies the (homogeneously) agent-neutral with the moral.

independence sense of the distinction, agent-relative reasons are dependent on the agent in a similar sense to that in which perceptual impression is dependent on the subject: without the subject/agent there are no such things. Agent-neutral reasons, on the other hand, on this dependence/independence sense, are more like objects than the perceptual impressions of objects; they exist independently, regardless of whether we have a perspective on them or not. Take definition A, above, as the standard definition of the agent-relative/agent-neutral distinction on this dependence/independence sense. This follows because, if – as it is on the dependence/ independence sense – an agent-relative reason is dependent on the agent, then that reason cannot be formulated, as definition A goes, without essential reference to the agent. And if – as it is according to the dependence/independence sense – an agent-neutral reason obtains independently of the agent, then it can be formulated without essential reference to the agent.

A primary example of the agent-relative/agent-neutral distinction in the dependence/independence sense is the metaethical contrast between Constructivism and Realism. According to the Constructivist, normativity is imposed upon the world by agents, and there would be no value in the world if we humans did not deem some things valuable. According to one foremost Constructivist account (Korsgaard 1996), an agent has a reason when he/she identifies him-/ herself in a particular way – say, as a violinist. Being a violinist gives such an agent a reason to practise playing the violin. This is dependence agent-relativity, because the counterfactual holds that, had the agent not identified him-/herself in that way, he/she would not have had such a reason – indeed, there wouldn't have *been* any such reason to be had at all. Had the agent not identified himself as a violinist, he would not have had a reason to practise playing the violin. On the other hand, according to Realism, normativity is not created by the agent, but is discovered, and there would be (would continue to be) value in the world if we humans were no longer around to recognise it. This is independence agent-neutrality, because the counterfactual does *not* hold that, had the agent not recognised the reason to Φ, he/she would not have had such a reason; rather, on this independently agent-neutral, Realist account, if the agent fails to recognise his/her reason to Φ, then the agent is simply mistaken.

The argument
So, because there are these two senses of the distinction – hetero-/homogeneous and dependence/ independence – and because these senses are entirely separate and distinct, but are nevertheless picked out by the self-same terminology of 'agent-relative' and 'agent-neutral', there is large scope for serious confusion. [11] The traditional, unqualified terminology ought therefore to be amended.

---

[11] That there are two separate definitions demarcated by the self-same terminology is, I think, sufficient to motivate the claim that there is large scope for serious confusion (and that, therefore, amendments ought to be made). However, my claim that the terminology is confusing does receive further support in later sections of the paper, in particular where, by way of application to Utilitarianism, I explore just how difficult it is to apply these labels of 'agent-relative' and 'agentneutral' even after we have distinguished between the two senses thereof.

The reduction objection

However, you might complain, what I have described as two separate and distinct senses of the agent-relative/agent-neutral distinction are really not that separate and distinct. Instead, you might suggest, definitions A and B are really talking about the same thing – just with B one step further removed than A, so to speak. That is, you might think that the hetero-/homogeneous sense in fact *reduces to* the dependence/independence sense. Here's how: if reasons are independently agent-neutral – that is, if they exist independently of our taking a view of/on them, like, say, this pen – neutral – that is, if they exist independently of our taking a view of/on them, like, say, this pen – then it follows that they will be agent-neutral in the homogeneous sense too. This pen is this pen is this pen, independently of what any of us think of it, so the pen's existence goes for *all* of us – that is, it is homogeneously agent-neutral. The independent agent-neutrality of the pen entails its homogeneous agent-neutrality. What applies in the case of pens applies in the case of reasons: if reasons are independently agent-neutral, then their existence goes for *all* of us, and they are thus homogeneously agent-neutral too. The independent agent-neutrality of reasons entails their homogeneous agent-neutrality.

If, on the other hand, reasons are dependently agent-relative – that is, if they are more like a perception of a pen than the pen itself – then it follows that they are heterogeneously agent-relative too. Consider how our perceptions of this pen are all slightly different: I see it from this angle, you see it from that angle; you see it head-on, another sees it length-ways. So, if reasons are agent-relative in the dependence sense, like perceptual impressions, then, as each person has a different perceptual impression, each person will have a different reason, and dependently agent-relative reasons are thus agent-relative in the heterogeneous sense too.

So, agent-relativity in the dependence sense entails agent-relativity in the heterogeneous sense, and agent-neutrality in the independence sense entails agent-neutrality in the homogeneous sense. And, the objection goes, the hetero-/homogeneous sense of the distinction is really just a derivative sense of the dependence/independence sense. At the end of the day, my two senses reduce to the same distinction, and I have no grounds for complaint.

Response

However, this objection depends on a misunderstanding of dependence agent-relativity and/or a lack of creativity. In response to the reduction objection, I argue that the hetero-/homogeneous sense of the distinction cannot simply be reduced to the dependence/independence sense, due to what I call 'combination cases', that is, cases of seriously plausible metaethics that are agent-relative in one sense yet agent-neutral in another. There are two kinds of combination cases available: metaethics according to which reasons are dependently agent-relative and nevertheless homogeneously agent-neutral, on the one hand, and metaethics according to which reasons are independently agent-neutral but nevertheless heterogeneously agent-relative. Of course, the more combination cases I can provide, the greater the ammunition for my critique of the traditional agent-relative/agent-neutral distinction. I think the former

kind of combination case is easily achieved, as soon as dependence agent-relativity is correctly understood. The latter kind of combination case, however, is not as straightforward, and to that end I examine three candidate cases below.

Combination case 1: dependence agent-relativity with homogeneous agent-neutrality

Contrary to the reduction objection, dependence agent-relativity *doesn't* actually entail heterogeneous agent-relativity, and we were misled into thinking that it does by a false analogy of dependence agent-relativity with perceptual impression. Agent-relativity in the dependence sense is *not* like perception – if this means: perception *of a thing that has independent existence*. The *point* of dependence agent-relativity is that, unlike independent agent-neutrality, *there's nothing more to reasons beyond the agent's relationship to them*. This is unlike perception, which (as opposed to hallucination) is always *of* something independent. For instance, consider Constructivism as a dependently agent-relative metaethic.[12] On this account the agent *gives him-/herself reasons* by way of his/her identification, e.g. an agent has reason to practise playing violin because he identifies himself as a violinist. This is *not* to say – as the pen analogy might have suggested – that the agent's identifying himself as a violinist (and thereby giving himself reason to practise playing violin) is a response to the perceived, independently-existing value of violin-playing – because, plausibly, violin-playing has no such independently-existing value. Rather, says Constructivism, violin-playing is valuable *because the agent values it* (only); we *give* it value, as we give ourselves reasons. Dependently agent-relative reasons are not perceptions of independently-existing value; perceptual impressions agent-relative reasons are not perceptions of independently-existing value; perceptual impressions of pens, on the other hand, are perceptions of independently-existing pens, and that is why the dependence agent-relativity of perceptual impression would seem to entail its heterogeneous agent-relativity: because we all occupy a different perspective on the independently-existing pen.

The pen analogy was originally employed to illustrate, by analogy with perception, the mind-dependent nature of agent-relativity (in the dependence sense) only. The pen analogy is a nice one, because it allows one to draw an easy comparison between the mind-dependent perception and the mind-independent object of perception, as we indeed did, above; but it's also misleading for this very reason. Mind-dependence is a feature of all mental states, which include – but are not limited to – perceptual impression, so we may well have used any other mental state besides perceptual impression to illustrate the mind-dependence of agent-relative reasons (in the dependence sense). So, rather than perceptual impressions, dependently agent-relative reasons might be more like, say, decisions: they're mind-dependent (as are all mental states, including decisions), but they're not responsive to the world in the tight kind of way that perceptual impressions are. Unlike perceptual impression, where we're all taking a (slightly) different perspective on an independently existing object, there's

---

[12] I gave a brief description of Constructivism at the end of the section titled 'The dependence/independence sense', above.

nothing in the nature of a *decision* that rules out the possibility of two people reaching the same decision; indeed, it's a common- place occurrence. Two people's reaching the same decision is possible *even if* those two people occupy radically different perspectives;[13] for instance, a middle-class, 30-something, stay-at- home mother might find the time to pursue her dream and finally decide to take up the violin, just as a 16-year-old, musically gifted pianist might want to increase his repertoire and decide to take up the violin. So, although it seems that, in the case of perceptual impression, dependence agent-relativity entails heterogeneous agent-relativity, the entailment need not hold in the case of a different mental state, like making a decision. There don't appear to be grounds for thinking that dependence agent-relativity entails heterogeneous agent-relativity when it comes to decisions. If dependently agent-relative reasons are, then, more like decisions and less like perceptual impressions, there is room for a metaethic according to which reasons are dependently agent-relative but nevertheless homogeneously agent-neutral.

However, is the mere *possibility* of dependently agent-relative, homogeneously agent-neutral reasons sufficient to support my argument? If I want to make the bold claim, as I do, that the traditional 'agent-relative'/'agent-neutral' terminology ought to be amended, then the combination case I wish to accommodate by that amendment must at least be a seriously plausible metaethic.[14] However, you may think, the combination case I've described above isn't plausible at all. In particular, you may think that it's highly *unlikely* that dependently agent-relative reasons would ever turn out to be homogeneously agent-neutral if they are less like perceptual impressions and more like decisions. (How often do different people reach the same decisions – such as taking up the violin, for instance – anyway?) Indeed, it seems that if dependently agent-relative reasons are, as I have urged, less like perceptions and more like decisions, then the chances that we would all have the same dependently agent-relative reasons would *decrease* dramatically. Why? Consider why we might expect people to share largely the same mind-dependent states. Take, again, the case of perceptual impression, for instance. We'd expect people to (largely) share the same perceptual impression when there is something out there that causes that perceptual impression; the pen, for instance. That is, large-scale homogeneity, if it is to be statistically likely and more than pure fluke, requires some kind of causal connection, some common cause that therefore explains the homogeneous effects. In the perceptual example, this role is served by the

---

[13] Objectors may complain that this is an illegitimate, equivocational use of 'perspective' in that it means something like a person's demographic here, whereas the previous use of the term meant something quite different: a person's direction of gaze, given their position in space. Objectors would be right to point out this difference. However, I deny that it is illegitimate: the point of my response is that dependently agent-relative reasons are not as closely analogous to perceptual impressions as the reduction objection supposes (and, by way of the new analogy proposed in its stead, that decisions are not closely analogous to perceptual impression either), and that therefore we shouldn't expect to see a neat corollary of perceptual-perspective in the realm of decisions. Whether or not the illustration I go on to provide is a legitimate case of like decisions produced by divergent perspectives, I think the point stands that there is nothing in the nature of decision that entails that different deciders must necessarily reach different decisions, where there apparently is in the case of perception.

[14] Bear in mind that 'seriously plausible' needn't mean *correct*. Insofar as the agent-relative/agent-neutral distinction is supposed to help settle metaethical debates (only), the combination cases I propose in favour of the amendment of that traditional distinction needn't be correct; they just need to be decent enough candidates.

pen. However, if dependently agent-relative reasons are less like perceptions and more like, say, decisions, when – as I put it above – the point of dependence agent-relativity is that there's nothing more to reasons beyond the agent's relationship to them, then any possible connection between the homogeneous mind-dependent states and their common cause is severed. Without anything independent to explain the homogeneous effects, it seems very unlikely indeed that we'd all end up with the same mind-dependent states. So, if reasons are dependently agent-relative, it's strikingly unlikely that they could be homogeneously agent-neutral. Therefore, the objection goes, we still have no serious cause to amend the traditional, very well established vocabulary of the 'agent-relative'/'agent- neutral', since the first proposed combination case really isn't seriously plausible.

However, this is short-sighted. I will grant to the objector that only a seriously plausibly metaethic could give us cause to amend the traditional agent-relative/agent-neutral distinction, and that a metaethic, the conditions of which are statistically unlikely to obtain, is not seriously plausible. I will even grant that homogeneity is unlikely to obtain when there is no common cause that would explain the homogeneity. However, what I dispute is that this common cause must be independent of the agent – something like a pen, for instance, out there in the world. Instead, I contend that there *is* something common amongst us, some shared feature that would increase the likelihood that we all end up having the same dependently agent-relative reasons. However, that shared feature isn't independent of us, it isn't out there in the world; rather, it is just this: the fact that we are all *agents*. It strikes me as quite plausible that there could be *something about us*, some feature of being human – not something in the world apart from us – that makes it quite likely, perhaps even necessitates, that we would all make the same decisions on some matters, even given the large variation of personality, age, culture, perspective, etc., amongst us. Indeed, I think this gets to the heart of Kantian metaethics. A metaethic that were, *à la* Kant, to posit dependently agent-relative reasons grounded not in the world but in our nature as agents, in our *humanity*, would have good cause to expect that those dependently agent- relative reasons would nevertheless turn out to be homogeneously agent-neutral, given that common cause that is our human nature.

So, there is at least one seriously plausible metaethic according to which reasons are dependently agent-relative, more like decisions than like perceptions, but which would nevertheless turn out to be homogeneously agent-neutral, given the way they are grounded in human nature, which is common to all of us. This neo-Kantian combination of the dependently agent-relative with the homogeneously agent-neutral gives us good cause to amend the traditional, unqualified distinction.[15]

---

[15] See Korsgaard (1993) for an account of this kind of neo-Kantian combination case, and for a critique of the traditional agent-relative/agent-neutral distinction similar to the one I have offered in this section.

Combination case(s) 2: independence agent-neutrality with heterogeneous agent-relativity

It may be possible to secure a second nail in the coffin for the traditional agent-relative/agent- neutral distinction by demonstrating a second kind of combination case. Whereas, above, it was argued that a certain kind of neo-Kantian metaethic can hold reasons to be dependently agent- relative but nevertheless homogeneously agent-neutral, this second combination case must be one according to which reasons are independently agent-neutral but nevertheless heterogeneously agent-relative. To this end, I examine below, three candidates for this second kind of combi- nation case: a modified version of Divine Command Theory, a modified version of Realism, and Utilitarianism.

Divine Command, to some but not to others, Theory

Imagine God tells some people, say, the Jews, not to eat certain things, but has nothing to say to the rest of the population on this score. As a result, some people have a reason to abstain from the rest of the population on this score. As a result, some people have a reason to abstain from certain foods, but others do not. That is, we have heterogeneous agent-relativity. Furthermore, since the reason came not from the agents but from *God*, we have independent agent-neutrality.[16]

Since many people do indeed believe that God commanded the Jews to behave in ways he didn't command the Gentiles to, I take this to be a seriously plausible metaethic. Thus, this case of the independently agent-neutral in conjunction with the heterogeneously agent-relative looks like a second case against the traditional, unqualified agent-relative/agent-neutral distinction.

However, I'm afraid, I think this is wrong. It is correct that a command given by God to some and not others would be heterogeneously agent-relative, but I do *not* think it correct that such a command would be *independently agent-neutral*. This is because such a reason would not issue from God *per se* but from *God's commanding*. While God *per se* is independently agent-neutral, God's commanding is not. This is because commanding, like any speech act, is a *second-personal* affair, as Stephen Darwall (2006) has pointed out. Commanding presupposes a listener as well as a speaker; there can be no such thing as a command unless there is, conceivably at least, *someone to be commanded*. So, says Darwall, a second-personal reason – like the reason issued by a command – is not (independently) agent-neutral but is, rather, (dependently) agent-*relative* (2006: p. 8).[17] Unlike God *per se*, God's commanding would depend as much on the Jews as it would on God himself, and is therefore a dependently agent-*relative* matter, rather than an independently agent- neutral one, as we originally supposed. So, God's commanding the Jews but not the Gentiles to abstain from certain foods would constitute not an independently agent-neutral but nevertheless heterogeneously agent-relative reason, as

---

[16] Thanks to Prof. Thaddeus Metz for this suggestion.

[17] Note that Darwall (2006) does not distinguish between the two senses of the traditional agent-relative/agent-neutral distinction as I have. However, that he gives 'agent-neutral' reasons the alternative title of 'state-of-world-regarding' reasons (p. 6) I think makes quite clear that he has in mind the dependence/independence sense of the distinction.

suggested above, but rather just a dependently agent- relative and heterogeneously agent-relative reason. Despite initial appearances, this version of Divine Command Theory does not constitute a combination case.

Conditional Realism

Since a version of Constructivism (namely, neo-Kantianism) afforded us the first combination case, according to which reasons are dependently agent-relative yet homogeneously agent-neutral, perhaps Realism – which I earlier contrasted with Constructivism – will afford us the second combination case, according to which reasons are independently agent-neutral yet heterogeneously agent-relative.

According to Realism, normativity is not created but discovered, and there would be (would continue to be) value[18] in the world if we humans were no longer around to recognise it. Realism is therefore an agent-neutral metaethic in the independence sense. But how can this be consistent with heterogeneous agent-relativity? Above, I argued that 'This pen is this pen is this pen, independently of what any of us think of it, so the pen's existence goes for *all* of us'; that is, the pen's independence agent-neutrality entails its homogeneous agent-neutrality. However, even if reasons are mind-independently real things, like pens, they needn't be real *in the same sense* as pens are real. Of course, reasons aren't *physically* real, as pens are. However, they may be real in a different kind of way. Rather, claim Realists, reasons are *sui generis* entities, independently agent-neutral, like pens, yet non-physical and normative, unlike pens. It is the Realist *sui generis* nature of reasons that opens the door for heterogeneous agent-relativity.

I have in mind the proposal that, if reasons are *sui generis* entities, these *sui generis* entities may have a *conditional* form. (If they're non-physical entities, what's stopping them from having a conditional form?) That is, they may have the form 'If you are such-and-such or are in such-and-such a position, then you ought to...'. I think it's quite plausible that reasons *do* have this conditional form. Consider an example from earlier: if there is a child drowning nearby, and if you conditional form. Consider an example from earlier: if there is a child drowning nearby, and if you can swim, then you have a reason to help save that child. That means that: if there *isn't* a drowning child, or if the drowning child *isn't* nearby, or if you *can't* swim (or all of the above), then you *don't* have a reason to try to save the child – *even when*, if there was a drowning child and she was nearby, somebody who could swim *would* have a reason to try to save the child. That is, reasons of this conditional form would be heterogeneously agent-relative, applying to those who fulfil the antecedent of the conditional but not to those who don't. Thus, we would have a metaethic that is independently agent-neutral, being a version of Realism, yet simultaneously heterogeneously agent-relative, given the conditional nature of the *sui generis* real entities. That is, we would have a second combination case against the traditional agent-relative/agent-neutral distinction.

---

[18] Strictly speaking, this claim ought to be restricted to 'moral value'. (The same goes for my mention of 'reasons' in this present section.) However, I have not made much of the moral/non-moral distinction in this paper. If a combination case can be established just with respect to moral reasons (but not with respect to non-moral reasons), that is sufficient for my argument.

However, is Conditional Realism a seriously plausible metaethic? As far as I am aware, no philosopher has yet defended a metaethic like it.[19] To this extent, it doesn't look as if Conditional Realism is seriously plausible. Traditional Realism, however, certainly is a seriously plausible and popular metaethic – and Conditional Realism does not strike me as too great an adaptation of traditional Realism. Furthermore, I would go so far as to claim that perhaps no philosopher has yet defended a metaethic such as Conditional Realism *because* the traditional, unqualified vocabulary has obscured its possibility. However, even if the reader is not prepared to accept an as-yet undefended metaethic as a seriously plausible combination case against the traditional distinction, according to which reasons are independently agent-neutral but nevertheless heterogeneously agent-relative, the following combination case will satisfy.

Utilitarianism

Utilitarianism is the flagship agent-neutral ethical theory. Any undergraduate Ethics student can tell you that Utilitarianism is agent-neutral. However, in what sense? It must be, it seems: agent-neutral in the sense that it doesn't matter *whose* utility is maximised by an action – your wife's, or the other drowning person's, for instance. As I described homogeneous agent-neutrality, above: 'it doesn't matter if you're the president of the universe or just the milkman'.[20] Therefore, Utilitarianism must be agent-neutral in the homogeneous sense.

However, this would be a mistake. The rest of my description of homogeneous agent-neutrality was: 'it doesn't matter if *you're* the president of the universe or just the milkman: *you have a reason* to jump into that pool and save the drowning child' (emphasis added).[21] That is, 'it doesn't matter' who the person is, as far as homogeneous agent-neutrality is concerned, when we're talking about the identity of the *agent*, the person who has the reason. (This is because, if reasons are homogeneously agent-neutral, *all* agents will be subject to the same reasons. So it doesn't matter who you are.) However, this is quite different to the present Utilitarian case. In this case, when we say that 'it doesn't matter' whether the drowning person is your wife or a stranger, we're talking about the identity of the *beneficiary*, the one to whom the action will be done – not the *agent*, the one who will act. This is a slippery point, but I think it's crucial. That it doesn't matter whose utility is maximised by an action is not what makes Utilitarianism agent-neutral.

This is not to say, however, that I disagree with thousands of Ethics undergraduates and the metaethical canon that Utilitarianism is agent-neutral. However, I argue, Utilitarianism is agent- neutral in the *independence* sense. However, if Utilitarianism is the thesis that we ought to maximise happiness – and happiness is surely mind-*dependent* – how could Utilitarianism be independently agent-neutral? I have a few points in response to this reaction. First, recall that Utilitarianism, generally speaking, is not

---

[19] Ronald N. Giere (2010), however, has defended a position called 'Conditional Realism' in the philosophy of science. However, this is apparently unrelated to the metaethical position I have in mind.
[20] This is from the section titled 'The hetero-/homogeneous sense'.
[21] This is from the section titled 'The hetero-/homogeneous sense'.

committed to the identification of utility with happiness in particular. Second, consider how Utilitarian reasoning is employed with regard to animals (e.g. Bentham and Second, consider how Utilitarian reasoning is employed with regard to animals (e.g. Bentham[22] and Singer[23]) or to future persons (e.g. Benatar[24]). Given that such beneficiaries aren't *agents*, the utility appealed to can't be dependently *agent*-relative, and so must be independently agent-neutral. Last, recall that Utilitarianism locates normativity not in utility *per se* but in the *maximisation* thereof. The Utilitarian principle calls us to '*maximise* expected utility', or, as it might otherwise be phrased, 'act so as to produce the greatest *net* utility or least *net* disutility'. Of course, an action that produces the least *net* disutility may actually produce some *utility*. If Utilitarianism located normativity in utility *per se*, such an action that produced a net disutility (some utility, but more disutility) would be partly good. However, this isn't the case. Utilitarianism locates normativity not in utility *per se* but rather in the *maximisation* thereof, such that an action that produces a net disutility is morally bad, even if that action produces some utility. Note that *maximisation* doesn't appear to be a mind-dependent matter. Maximisation is, rather, an objective fact, a state of the world (even if it is a state of the world-as-containing-minds-that-experience-utility and those objective facts are objective *psycho- logical* facts). [25] Utilitarians, I think, conceive of utility as something like, say, oxygen: oxygen can be in your lungs or in my lungs, but there is an objective matter of fact as to what atmospheric oxygen levels are (and we want those levels to be higher rather than lower). Seeing, on this analogy, how the Utilitarian thinks about utility can also give us insight into why it is, on the Utilitarian account, that 'it doesn't matter' whether it's your wife or the other drowning person whose utility is maximised: utility is just utility, as oxygen is just oxygen, even though sometimes a bit of it is in you and sometimes a bit of it is in me, or in your wife or in the other drowning person. Thus, the way in which utility and its maximisation is understood is as independently agent-neutral. This holds even if utility is identified with happiness.

Furthermore, I wish to argue that the sense in which Utilitarianism is agent-neutral is not the homogeneous sense (as we first considered) because Utilitarianism – if it is a seriously plausible metaethic – is actually *heterogeneously agent-relative*. If this is correct, and if my argument that Utilitarianism is independently agent-neutral (above) is correct, then Utilitarianism constitutes a combination case against the traditional distinction, a metaethic according to which reasons are independently agent-neutral but simultaneously heterogeneously agent-relative. [26] Why think that Utilitarianism is

---

[22] See *The Principles of Morals and Legislation* (1948).

[23] See *Animal Liberation* (1975).

[24] See *Better Never to Have Been* (2008). Thanks to Anna Hartford for helping me with this point.

[25] Thanks for Prof. Thaddeus Metz for help with this point

[26] Readers unsatisfied with my argument (above) that Utilitarianism is agent-neutral in the independence sense may wish to retain the initial impression that Utilitarianism is agent-neutral in the sense that it doesn't matter whose utility is maximized – your wife's, or the other drowning person's. I have shown, above, that this 'it doesn't matter whose' is not homogeneous agent-neutrality – so my critic may wish to argue, instead, that this demonstrates that there is a further, third sense of the agent-neutral. (Perhaps call it the love/beneficence sense of the distinction, following Harry Frankfurt's (2004) claim that love is 'ineluctably personal' in that the object of one's love cannot be substituted (p. 79) – that is, agent-relativity in this sense – whereas, when it comes to mere beneficence (Frankfurt gives the example of a charity worker), 'any sick or poor person will do' (p. 43) – that is, agent-neutrality in this sense.) I am quite

heterogeneously agent-relative? To see why, consider what things would look like if Utilitarianism were not heterogeneously agent-relative but homogeneously agent-neutral. Say, for instance, the action of saving a drowning child in Papua New Guinea is the action that, now, would produce the greatest net utility – greater than any other action I could now perform in South Africa. (We can imagine that the child is perhaps some sort of prodigy – say, a medical genius, who is on the brink of developing a cure for AIDS, that this child has captured the world's heart and that millions of people would be devastated by her death, etc.) According to Utilitarianism, then, I have a reason to try to save the drowning child in Papua New Guinea (as does anyone and everyone, if Utilitarianism is homogeneously agent-neutral). However, this isn't intuitively plausible. First, I'm not *in* Papua New Guinea; I'm here, on the other side of the world. So I'd have to get myself to Papua New Guinea. However, by the time I got to Papua New Guinea, some other agent, who – if Utilitarianism is homogeneously agent-neutral – would also have a reason to try save the drowning child, and who is a resident of Papua New Guinea, would probably already have saved the child. Thus, on this homogeneously agent-neutral account of Utilitarianism, I would have been morally obliged to fly myself to Papua New Guinea, since that was a means to the end of producing the greatest net utility, but the chances of this morally required action *actually producing* the greatest net utility – or *any* utility at all – are really quite small. Intuitively, I don't have an obligation to try to save the drowning child in Papua New Guinea, since it's highly unlikely that I could do anything to help her. However, those agents who are residents of Papua New Guinea, and who are in the vicinity of the drowning child, intuitively *do* have a reason to try to save her, since it is likely (or, at least, more likely than it is in my case) that their actions will be successful and will indeed maximise utility. If, intuitively, some agents will have a reason to try to save the drowning child, but others will not, this is an appeal to heterogeneous agent-relativity.

Indeed, one way to read the qualification of the Utilitarian principle that we ought to 'maximise *expected* utility' might be as an attempt to capture this heterogeneous agent-relativity. The 'expected' qualification, cashed out, is this: you have a reason to act so as to maximise utility *when you could reasonably expect that that action will indeed maximise utility.*[27] Thus, conversely, when you can't reasonably expect that your action will indeed maximise utility, you don't have a reason to so act. When, for instance, it is highly unlikely that my trying to save the drowning child in Papua New Guinea will

---

sympathetic to this. It would constitute a critique of my characterisation of Utilitarianism – but it would nevertheless constitute a combination case to challenge the traditional agent-relative/agent-neutral distinction, in that Utilitarianism would be agent-neutral (in the beneficence sense) yet agent-relative (in the heterogeneous sense, as I go on to argue).

[27] I say in the opening sentence of this paragraph that this is 'one way' to understand the 'expected' qualification, since I think that the standard way of reading it is different to the reading I present above. The standard reading of this qualification, cashed out, focuses on the 'could reasonably expect' component, such that the converse of it is: when unusual, idiopathic preferences mean that Φing will produce a very large amount of utility – but you could not have known that Φing would have had that consequence – then you did not have a reason to Φ, and you are not blameworthy for not Φing. That is, the standard reading of the 'expected' qualification is an epistemological one: you don't have a reason when you couldn't have known the consequences. The reading I offer here, however, is a practical reading. It focuses on the 'will indeed maximise utility' component, such that the converse is: you don't have a reason to Φ when your Φing won't actually produce the desired consequences. These two readings, the epistemological and

indeed be successful and will maximise utility, I don't have a reason to try to save her. However, when, on the other hand, we can reasonably expect that the action of a resident of Papua New Guinea will be successful and will maximise utility, then that agent *does* have a reason to try to save the drowning child. Heterogeneous agent-relativity, it seems, is built into the Utilitarian principle that we ought to maximise *expected* utility.[28]

I do not think that any of this discussion constitutes a significant variation on traditional Utilitarianism – unlike Conditional Realism with regard to traditional Realism – so, as such, I do not think the question of Utilitarianism's being a seriously plausible metaethic needs to be addressed. Its popularity proves as much. I therefore conclude that Utilitarianism, well-known as it is for being 'agent-neutral', is agent-neutral only in the independence sense, yet is simultaneously agent-*relative* in the heterogeneous sense, and therefore constitutes a further combination case against the traditional agent-relative/agent-neutral distinction.

Conclusion

In this paper, I have delineated two senses of the widely used agent-relative/agent-neutral distinction: the hetero-/homogeneous sense (according to which agent-relative reasons differ from person to person, and agent-neutral reasons obtain uniformly across persons) and the dependence/independence sense (according to which agent-relative reasons depend essentially on the agent, whereas agent-neutral reasons do not). I have defended my claim against the objection that the hetero-/homogeneous sense ultimately bottoms out as the dependence/independence sense by demonstrating that we can have combinations of agent-relativity in the one sense with agent-neutrality in the other. The first combination case against the traditional distinction was that of a neo-Kantian metaethic, according to which reasons are agent-relative in the depend- ence sense, but turn out to be agent-neutral in the homogeneous sense, grounded as they are in our humanity, which is necessarily a common feature of all of us. A further combination case against the traditional distinction – a case of independently agent-neutral reasons which are nevertheless heterogeneously agent-relative – was found in Utilitarianism, where I argued that Utilitarians conceive of utility

---

the practical, are nevertheless consistent with each other. I suggest that Utilitarians have in mind both with their qualification that we ought to act so as to maximise expected utility.

[28] My thanks to Prof. Thaddeus Metz and this journal's anonymous reviewer for helping me to see this point. In the original presentation of this paper, I argued that Utilitarianism, like Divine Command Theory, fails to constitute a combination case. My original argument was as follows. Utilitarianism is the fl agship agent-neutral metaethic – where this is agentneutrality in the independence sense (as above). Then, I argued that, in order for Utilitarianism to cope with the Papua New Guinea case, Utilitarianism must adopt a 'probability constraint' according to which an agent S has a reason to Φ when: (1) Φing maximises utility, and (2) the 'probability constraint' is met: it is reasonably likely that S's Φing will maximise utility. However, if Utilitarianism adopts this probability constraint (which it must, I argued, in order to cope with the Papua New Guinea case), then it in fact doesn't constitute a combination case, since the addition of criterion (2) transforms this metaethic from an independently agent-neutral one to a dependently agent-relative one, since the normative force on this probability-constraint version derives from the agent's being in the most likely position to maximize utility – that is, a feature of the agent, which is dependence agent-relativity and not the requisite independence agentneutrality. Prof. Metz has helped me to see that what I called the probability constraint may actually be built into the 'expected' qualification of the Utilitarian principle, and my argument has since changed accordingly.

(specifically, the maximisation thereof) as independently agent-neutral, and that the Utilitarian qualification to 'maximise *expected* utility' entails heterogeneous agent-relativity. Generous readers may, furthermore, be willing to concede that Conditional Realism constitutes a third combination case, according to which reasons are independently agent-neutral, real and *sui generis*, yet simultaneously heterogeneously agent-relative, given their conditional form. Each of the combination cases demonstrates, *contra* the reduction objection, that the hetero-/ homogeneous sense isn't simply a derivative of the dependence/independence sense. Given that there are (at least) two senses of the traditional 'agent-relative'/'agent-neutral' vocabulary at large in the literature, each future use of those terms ought to be stipulated as being used in either the dependence/independence or the hetero-/homogeneous sense, or the traditional terminology ought to be abandoned entirely, in favour of new terminology less likely to lead to confusion.

References

Darwall, S. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, Massachusetts: Harvard University Press.

Frankfurt, H.G. 2004. *The Reasons of Love*. Princeton: Princeton University Press.

Giere, R.N. 2010. 'Contingency, conditional realism, and the evolution of the sciences', *PhilSci- Archive*, available from http://philsci-archive.pitt.edu/8365/ [accessed 7 February 2013].

Korsgaard, C.M. 1993. 'The reasons we can share: an attack on the distinction between agent-relative and agent-neutral values', available from http://nrs.harvard.edu/urn-3:HUL.InstRepos:3196321 [accessed 17 September 2011].

Korsgaard, C.M. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press. Nagel, T. 1970. *The Possibility of Altruism*. Oxford: Oxford University Press.

Nagel, T. 1986. *The View From Nowhere*. New York: Oxford University Press.

Ridge, M. 2005. 'Reasons for action: agent-neutral vs. agent-relative', in: Zalta, E.N. (ed.). *Stanford Encyclopedia of Philosophy*. Stanford: Stanford University, available from http://plato.stanford.edu/entries/reasons-agent/ [accessed 23 October 2011].