

Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection School of Social Sciences

School of Social Sciences

5-2018

Distinguishing agent-relativity from agent neutrality

Matthew HAMMERTON

Singapore Management University, mhammerton@smu.edu.sg

Follow this and additional works at: https://ink.library.smu.edu.sg/soass_research



Part of the [Sociology Commons](#), and the [Theory and Philosophy Commons](#)

Citation

HAMMERTON, Matthew.(2018). Distinguishing agent-relativity from agent neutrality. *Australasian Journal of Philosophy*, , 1-12.

Available at: https://ink.library.smu.edu.sg/soass_research/2504

This Journal Article is brought to you for free and open access by the School of Social Sciences at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School of Social Sciences by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email library@smu.edu.sg.

Distinguishing Agent-Relativity from Agent-Neutrality

Matthew Hammerton 

Singapore Management University

ABSTRACT

The agent-relative/agent-neutral distinction is one of the most important in contemporary moral theory. Yet providing an adequate formal account of it has proven to be difficult. In this article I defend a new formal account of the distinction, one that avoids various problems faced by other accounts. My account is based on an influential account of the distinction developed by McNaughton and Rawling. I argue that their approach is on the right track but that it succumbs to two serious objections. I then show how to formulate a new account that follows the key insights of McNaughton and Rawling's approach yet avoids the two objections.

ARTICLE HISTORY Received 8 July 2017; Revised 19 April 2018

KEYWORDS agent-neutral; agent-relative; moral rules; ensuring; maximizing

1. Introduction

The distinction between moral rules, reasons, or theories that are agent-relative and those that are agent-neutral is one of the most important in normative ethics.¹ It is essential for understanding debates around deontic constraints, special duties, personal prerogatives, and consequentializing moral theories. Furthermore, several moral philosophers regard it as the central fault line in moral theory, marking two rival approaches to ethics.² Because of its importance, many have attempted to give a formal account of the distinction—appealing to the underlying logical structure of moral rules to mark them (and the theories that contain them) as relative or as neutral. The most influential and promising approach in this tradition is that developed by McNaughton and Rawling [1991, 1992, 1993, 1995a, 1995b, 1998, 2003]. In this article I raise two objections against their account. I then suggest a new way of marking the agent-relative/agent-neutral distinction that adopts the basic strategy of McNaughton and Rawling's approach, but that avoids the two objections. I argue that this is the best formal account of the distinction.

2. The Intuitive Distinction

The key idea behind the agent-relative/agent-neutral distinction is the thought that some moral rules (or reasons, or theories) give agents a special focus on themselves, their actions, or states of affairs concerning them, whereas other rules lack this personal

¹ Thomas Hurka [2003: 628] describes it as one of 'the greatest contributions of recent ethics'. Ridge [2005] and Portmore [2013] express similar views.

² For different views of why it is the central fault line in ethics, see McNaughton and Rawling [1992], Dreier [1993], and Pettit [1997].

focus and instead give agents a general concern with actions and states of affairs that may involve anyone. Parfit [1979, 1984] develops this idea intuitively by suggesting that agent-relative rules give different ultimate aims to different agents, whereas agent-neutral rules give to all agents the same ultimate aims. To see how Parfit's suggestion works, consider the following ten moral rules:

- (1) Each agent must not kill innocent people.
- (2) Each agent must minimize the killing of innocent people.
- (3) Each agent must care for her children.
- (4) Each agent must ensure that children are cared for by their parents.
- (5) Each agent must maximize her own well-being.
- (6) Each agent must maximize general well-being.
- (7) Each agent must maximize the well-being of her friends.
- (8) Each agent must ensure that no one is disloyal to the Supreme Leader.
- (9) Each agent must minimize her violations of rule (1).
- (10) Each agent must minimize general violations of rule (1).

Moral philosophers discussing the agent-relative/agent-neutral distinction have generally agreed that rules resembling the odd-numbered rules above (1, 3, 5, 7, 9) are agent-relative, whereas those resembling the even-numbered rules (2, 4, 6, 8, 10) are agent-neutral.³ Let's consider how Parfit's intuitive account explains these classifications.

Rule (1) gives each agent a special concern with her own acts of killing, requiring that she does not kill an innocent person even if it is the only way to prevent more killings of innocent people being performed by others. In this way, (1) gives different aims to different agents. For example, it gives Fred the aim that 'Fred does not kill innocent people' and Hema the different aim that 'Hema does not kill innocent people.' By contrast, rule (2) gives each agent a general concern with acts of killing committed by anyone, rather than a special concern with her own acts of killing. It requires an agent to kill an innocent person in any case where doing so is the only way to prevent more killings of innocent people being committed by others. Given this result, (2) appears to give all agents the same aim—the aim that the killing of innocent people is minimized.

Similar reasoning can be applied to the other rules. Rule (3) is agent-relative because it gives each agent a special concern with *her* own acts of caring, giving her the aim that she cares for her children. By contrast, (4) is agent-neutral because it gives all agents the same aim—that all children are cared for by their parents. Likewise, (5) is agent-relative because it gives each agent the aim that *her* own well-being is maximized, whereas (6) is agent-neutral because it gives all agents the same aim—that everyone's well-being is maximized. Rule (7) is agent-relative because it gives each agent the different aim that the well-being of *her* friends is maximized, whereas (8) is agent-neutral because it gives all agents the same aim—that no one is disloyal to the Supreme Leader. Finally, (9) is agent-relative because it gives each agent a special concern with her own violations of (1), whereas (10) is agent-neutral because it gives all agents the common aim that all violations of (1) are minimized.

³ For discussion of the classification of rules like those considered above, see Nagel [1970], Parfit [1984], Nagel [1986], McNaughton and Rawling [1991, 1992], Ridge [2005], and Portmore [2013].

Parfit's intuitive account has been very influential, as it appears to elegantly explain the standard classifications given to rules like (1)–(10). However, his account lacks the precision of a formal account, instead relying on our intuitive judgments about what ultimate aims a moral rule (or reason, or theory) gives to agents. Yet it is controversial whether all moral rules literally give aims to agents, leading to the concern that Parfit's talk of the 'ultimate aim' given by a moral rule is, at best, metaphorical [Portmore 2013]. Furthermore, some, such as Schroeder [2011: 36–9], are sceptical about whether there is a clear distinction here at all. For these reasons, a formal account of the agent-relative/agent-neutral distinction that explains the standard classification given to rules like (1)–(10) is highly desirable.

3. Developing a Formal Account

Formal accounts of the agent-relative/agent-neutral distinction have been developed by Nagel [1970], Sen [1982], Nagel [1986], Pettit [1987], McNaughton and Rawling [1991], and Skorupski [1995]. Of these various accounts, the account by McNaughton and Rawling is generally considered the most promising, and has come to be the most influential formal account in the literature.⁴ By contrast, the accounts by Nagel [1970], Sen [1982], Nagel [1986], Pettit [1987], and Skorupski [1995] are undermined by several strong objections that have been developed in the literature.⁵ For this reason, I will take McNaughton and Rawling's account as my starting point, attempting to show that it is on the right track and yet is undermined by two serious objections.

Let's start with an overview of their account. Their first step is to take all moral rules and transform them into semiformal statements with the following form: '(x)(x must ensure that [...]).' Thus, consider rules (1)–(6) above. McNaughton and Rawling's approach transforms these rules, as follows:

- (1*) (x)(x must ensure that [x does not kill innocent people])
- (2*) (x)(x must ensure that [(y)(y does not kill innocent people)])
- (3*) (x)(x must ensure that [(y)(y is x's child → x cares for y)])
- (4*) (x)(x must ensure that [(y)(z)(y is a child of z → z cares for y)])
- (5*) (x)(x must ensure that [x has maximal well-being])
- (6*) (x)(x must ensure that [(y)(y has maximal well-being)])

Their second step is to observe that, when rules are transformed in this way, all agent-relative rules have an occurrence of 'x' in the square brackets that is bound by the initial universal quantifier (for example, (1), (3), (5)), whereas all the agent-neutral rules have no such occurrence (for example, (2), (4), (6)). This observation leads to the following precise statement of the distinction:

A rule expressing a moral requirement is agent-relative *iff*, when the rule is represented with the form '(x)(x must ensure that [...])', there is an ineliminable occurrence of 'x' in the square brackets that is bound by the initial universal quantifier.

A rule expressing a moral requirement is agent-neutral *iff* it is not agent-relative.

⁴ For example, their original account [1991] is explored, endorsed, or made use of by Howard-Snyder [1996], Portmore [2001, 2013], Louise [2004], Dougherty [2013], Hammerton [2016], and McNaughton and Rawling [1992, 1993, 1995a, 1998, 2001].

⁵ For objections to these other accounts, see McNaughton and Rawling [1991, 1992, 1995b], Broome [1995], Ridge [2005], and Portmore [2013].

The main reason why McNaughton and Rawling's account has become the leading one is that it appears to correctly classify rules like (1)–(10) as agent-neutral or agent-relative, whereas rival accounts have trouble with this. For example, the accounts by Nagel [1986] and Pettit [1987] appear to misclassify (4) as agent-relative, whereas McNaughton and Rawling's account correctly classifies it as agent-neutral.⁶ A second attraction of their account is that the formal feature that they use to mark the distinction appears to correspond to the intuitive idea that some moral rules give agents a special focus on themselves, their actions, or states of affairs involving them, whereas others do not. For McNaughton and Rawling show how various moral rules can be translated into rules requiring agents to ensure that some state of affairs obtains, and they then tie agent-relativity/neutrality to the presence or absence in each rule's propositional content of a variable that refers to the agent who is bound by the rule. A final attraction of their account is that it is reasonably simple and elegant.

I will end this section with an important clarification of McNaughton and Rawling's use of the term 'ineliminable'. They include this term in their account because there is a trivial way in which any moral rule transformed into the canonical form can be made to have an occurrence of 'x' in the square brackets. For example, the agent-neutral (2) could be transformed as follows:

(2**) (x)(x must ensure that [x ensures that ((y)(y does not kill innocent people))])

McNaughton and Rawling point out that although (2) can be translated as (2**) we should not thereby regard (2) as agent-relative. This is because the occurrence of 'x' in the square brackets in (2**) is eliminable, as (2**) is equivalent to (2*) and there is no occurrence of 'x' in the square brackets of (2*). By contrast, the agent-relative (3) cannot be translated into the canonical form without the variable 'x' appearing in the square brackets, and thus the presence of 'x' in this rule's square brackets is ineliminable. This clarification is important because it is sometimes claimed that maximizing-rules like (2) ought to be translated into the canonical form as follows:

(2***) (x)(x must ensure that [x minimizes the killing of innocent people])

If (2) is translated this way, then it appears agent-relative as 'x' occurs in the square brackets. However, McNaughton and Rawling [1991: n7, 1993: 83–6] respond by pointing to eliminability. They claim that (2***) is not agent-relative, because it is equivalent to (2*), which demonstrates that the occurrence of 'x' in the square brackets in (2***) is eliminable. However, it is not clear to me (and others) that (2*) and (2***) really are equivalent. Below, I will develop an objection against McNaughton and Rawling's account that undermines their translation of (2) as (2*). I will then suggest an alternative translation for (2) that classifies it as agent-relative, and yet is clearly equivalent to sentences like (2***) that contain the relevant variable.

4. Two Objections against McNaughton and Rawling

McNaughton and Rawling offer a promising account of the agent-relative/agent-neutral distinction. However, there are two problems with their account that arise when they transform moral rules into the semiformal notation. As McNaughton and Rawling

⁶ See McNaughton and Rawling [1991: 171–2]

acknowledge, these transformations must preserve the content of the rules that are transformed.⁷ This is essential because, if the transformation of a rule changes its content, then showing that the transformed rule has a logical property that we associate with agent-relativity fails to explain why the original rule (with its different content) might count as agent-relative. I show in my objections that there are two ways in which McNaughton and Rawling's proposed transformations fail to give us accurate translations.

My first objection concerns the step from (1) to (1*). Recall (1):

(1) Each agent must not kill innocent people.

McNaughton and Rawling translate this rule into their canonical form, as follows:

(1*) (x)(x must ensure that [x does not kill innocent people])

If this translation is correct, then 'requiring an agent not to kill' must be equivalent to 'requiring an agent to ensure that she does not kill.' More generally, if direct-rules like (1) can be translated as ensuring-rules like (1*), then every rule requiring an agent to not perform a certain act must be equivalent to a corresponding rule requiring that agent to ensure that she does not perform that act. However, these things are not equivalent, as there are counterexamples where an agent does not perform an action without ensuring that she does not perform it. In these cases, the agent obeys a direct-rule without obeying the corresponding ensuring-rule. For example, suppose that Lee aims his rifle at the President's head and pulls the trigger with the intention and expectation that in doing so he will assassinate the President. Furthermore, suppose that, by chance, the President turns his head just as Lee fires, fortuitously moving it out of the path of the oncoming bullet and narrowly escaping the assassination attempt. In such a scenario, it is true that Lee has not done a certain act (the act of killing the President), yet it seems false that he has ensured that he did not do this act. After all, he has tried his best to make it the case that he *did* do this act, and it is odd to describe a circumstance where an agent earnestly tries to perform an act but fails due to bad luck as one where 'he ensures that he does not do the act.'

The above example is sufficient to prove that 'x does not φ ' and 'x ensures that x does not φ ' are not equivalent. Nonetheless, a further counterexample against the equivalence claim is also helpful. I have not performed the action of killing Napoléon Bonaparte. However, it is false that I have *ensured* that I have not killed him. Napoléon Bonaparte died in 1821. Therefore, no act or omission of mine has ever made it either more or less likely that I will not kill him.⁸

These examples show that McNaughton and Rawling's translation of (1) as (1*) fails. Further reflection shows that a rule like (1) cannot be transformed into the canonical form without altering its content. For there is no alternative to (1*) that both fits the canonical form and is an accurate translation of (1). Thus, McNaughton and Rawling's account of the agent-relative/agent-neutral distinction fails. To draw the distinction with formal precision, they assume that all moral requirements can be translated as rules with the form: '(x)(x must ensure that [...]).' However, a rule like (1) cannot be

⁷ McNaughton and Rawling [1993: 85] say, 'Transmutation to form F must be transmutation without alteration in ethical content.'

⁸ I thank Daniel Nolan for suggesting this example. It should be noted that, for the sake of argument, I am assuming in this example that time travel is nomologically impossible.

successfully translated in this way, and thus their account fails to classify (1) as agent-relative.

My second objection to McNaughton and Rawling's account concerns their translations of maximizing/minimizing rules. For example, recall rules (2), (5), and (6):

- (2) Each agent must minimize the killing of innocent people.
- (5) Each agent must maximize her own well-being.
- (6) Each agent must maximize general well-being.

McNaughton and Rawling transform these rules into the canonical form, as follows:

- (2*) (x)(x must ensure that [(y)(y does not kill innocent people)])
- (5*) (x)(x must ensure that [x has maximal well-being])
- (6*) (x)(x must ensure that [(y)(y has maximal well-being)])⁹

My objection is that (2*), (5*), and (6*) are not adequate translations of (2), (5), and (6). More fundamentally, my objection is that there is no satisfactory way to translate maximizing rules into McNaughton and Rawling's canonical 'ensure that' form. The objection can initially be pressed by pointing out that, in certain cases, rules like (2) and (2*) give different deontic verdicts, and thus cannot be equivalent. For example, consider the following case:

The Super Crime-Fighter. Clark Kent devotes his life to preventing the killing of innocent people, often performing superhuman feats and pushing himself to his limits. His efforts are a great success, and innocent people are rarely killed in his society. However, occasionally he finds himself unable to prevent all such killings. For instance, sometimes supervillains in different locations are simultaneously killing random innocent people and, although Clark can reach one location in time to prevent the killing, it is impossible for him to reach both. In these situations, he prioritizes the location where the most killings of innocents can be prevented, allowing the killings at the other location to occur.

In this example, Clark appears to obey rule (2). For, at each decision point that he faces, he always takes the available option that will result in the fewest killings of innocent people, and this is what it means to *minimize* such killings. However, Clark appears to disobey rule (2*). For, despite his tremendous efforts, innocent people are occasionally killed, and when this happens he has failed to ensure otherwise. Therefore, rules (2) and (2*) have different content.

McNaughton and Rawling might respond to this objection by pointing to a clarification they make [1995a: 34]. There, they state that 'x must ensure that ...' is shorthand for 'x must ensure, to the best of x's abilities, and in so far as there is no conflicting duty of greater weight, that ...'. Thus, their complete translation of (2) is this:

- (2[†]) (x)(x must ensure, to the best of x's ability, that [(y)(y does not kill innocent people)])

This translation may help them to deal with the Clark Kent example, as they can say that, although Clark Kent does not prevent every killing of an innocent person, he does prevent them to the best of his ability, and thus obeys (2[†]). More generally, they can say that ensuring something to the best of your ability is the same as maximizing that thing.

⁹ McNaughton and Rawling only explicitly consider rule (2), translating it as (2*). Rules (5*) and (6*) are my best attempt at fitting (5) and (6) into their framework.

However, Skorupski [1996: 240] has raised a problem with their use of the ‘best of x ’s ability’ clause. McNaughton and Rawling say that all rules are qualified with this clause. However, if that is the case, and this clause expresses the notion of maximization, then non-maximizing rules are wrongly translated by their account as being maximizing rules. For example, if (1) is qualified with the ‘best of x ’s ability’ clause, then (1) becomes a rule requiring each agent to minimize her killings of innocent people, making (1) equivalent to rule (9). Yet (1) is not a maximizing-rule. It is a rule requiring agents to never kill, classing them as doing the wrong thing any time that they kill, even if they are killing to minimize their total killings. Thus, McNaughton and Rawling appear to face a dilemma: their canonical form either fails to accurately translate maximizing-rules, or, with appropriate additional clauses, it successfully translates these rules but at the expense of mistranslating constraints like (1).

We can add to the above concern a second problem for McNaughton and Rawling’s translation of maximizing-rules—a problem related to the first objection above. It is possible for an agent to ‘accidentally’ maximize the non-killing of innocent people, by performing an action that results in non-killing being maximized without her intending or foreseeing that her action has this result. In such cases, although the agent maximizes, she does not appear to *ensure* that she maximizes. Therefore, it is possible to obey (2) and yet disobey (2[†]), and hence they are not equivalent.

5. A Better Formal Account

Above, we have seen that McNaughton and Rawling’s canonical ‘ensure that’ form is unable to adequately accommodate direct-rules and maximizing-rules. Some may be tempted to conclude from this that McNaughton and Rawling’s method for capturing the difference between agent-neutral and agent-relative rules is unworkable, and that a fundamentally different account of the distinction is needed. However, this is too hasty. For, although their canonical form is inadequate, one of their key ideas is very promising. They realized that, if all moral rules are expressible such that these contain propositional content, then the difference between agent-relative and agent-neutral rules might be discernible in that content.¹⁰ More precisely, McNaughton and Rawling noted that the difference may be marked by the presence or absence of a certain kind of variable in a rule’s content. Their mistake was to think that all rules can be expressed as ensuring-rules, when only some rules are expressible in this way. However, their basic strategy might succeed if alternative translations that contain the appropriate propositional content can be found. In this section, I will show how to translate direct-rules and maximizing-rules into semiformal notation, such that they contain propositional content, avoid the objections raised above, and have the appropriate formal properties to be classified correctly. This will lead to a new account of the agent-relative/agent-neutral distinction.

Let’s start with direct-rules. We saw above that ‘ x does not φ ’ and ‘ x ensures that x does not φ ’ are not equivalent, and so (1) cannot be translated as (1*). However, if we

¹⁰ It should be noted that, when I talk of moral rules containing propositional content, I am not taking a stand on whether ‘ought’ statements contain a relation between an agent and a proposition (as argued by Broome [1999] and Wedgewood [2006]), or an agent and an action (as argued by Schroeder [2011]). For an ‘ought’ statement may contain propositional content directly by taking a proposition as one of its relata, or indirectly by taking as one of its relata an action that contains propositional content (e.g. the action of ‘ensuring that p ’).

can find an alternative verb to ‘ensure’, one that both takes propositional content and can be used to generate an accurate translation of (1), then we can construct an account that accommodates direct-rules. Hence, we are looking for a rule with the form ‘ x {verb} that x does not φ ’ that is equivalent to ‘ x does not φ .’ One possibility arises from using the verb ‘making it the case’. Perhaps ‘ x makes it the case that x does not φ ’ is equivalent to ‘ x does not φ .’ To test this equivalence claim, let’s consider the two counterexamples from section 4. In the chancy-failed-assassination case, ‘make it the case that’ appears to do better than ‘ensure’. For, although Lee attempts to make it the case that he assassinates the President, it seems plausible to say that, by firing his bullet at that exact moment in that exact way, he is (by accident) making it the case that he does *not* assassinate the President at that moment. However, despite doing well on the first example, this verb gives the wrong verdict in the Napoléon example. For it is true that I do not kill Napoléon Bonaparte and yet false that I make it the case that I do not kill him. The latter is false because ‘making it the case’ suggests that I have done some act (or perhaps some omission) that causes the relevant state of affairs to obtain. Yet, given the physical impossibility of my killing Napoléon Bonaparte, it seems false to say that my acts (or omissions) have caused this negative state of affairs to obtain. It follows from this that direct-rules are not equivalent to their corresponding make-it-the-case-that rules.

Two other candidates are the verbs ‘bring it about that’ and ‘see to it that’.¹¹ Perhaps ‘ x brings it about that x does not φ ’ or ‘ x sees to it that x does not φ ’ are equivalent to ‘ x does not φ .’ However, because of the causal connotations that they carry, both of these verbs fail in the Napoléon example for the same reason that ‘make it the case that’ failed. I do not kill Napoléon Bonaparte, and yet it is false that I bring it about that, or see to it that, I do not kill him. Thus, we are looking for a verb that is appropriately connected to an agent not performing an act, without implying that the agent did something that caused herself not to perform that act.

I propose the verb ‘act such that’ as the solution to this problem. Unlike the other verbs that we have considered, it appears to avoid both counterexamples. In the chancy-failed-assassination case, it is true that Lee does not assassinate the President, and also true that he acts such that he does not assassinate the President. His firing of the bullet at that exact moment in that exact way is one way of acting such that he does not assassinate the President. Likewise, in the Napoléon example, it is true that I do not kill Napoléon Bonaparte, and also true that I act such that I do not kill him. This is because ‘acting such that’ does not carry the causal connotations that ‘making it the case that’ carries. Instead it means something weaker, along the lines of ‘performing acts or omissions that, given the state of the world, are consistent with it being the case that’.¹² And, at each moment when I do not kill Napoléon Bonaparte, it is also true that I am performing acts or omissions that, given the state of the world, are consistent with my not killing him. Of course, this is true in a trivial way because any act or

¹¹ ‘See to it that’ has pedigree in deontic logic. For example, see Belnap and Perloff [1988].

¹² This is the reading that I find most natural for the locution ‘act such that’. However, my argument does not require that this is the best reading of this locution. All that I need to show is that there is a possible verb (perhaps one that I define by stipulation) with the relevant properties.

omission that I might perform is consistent with my not killing him, but it is true nonetheless.

The results that we have just seen, together with reflection on further cases, show that ‘ x acts such that x does not φ ’ is true if, and only if, ‘ x does not φ ’, and thus they are equivalent. It follows that, although (1) cannot be accurately translated as (1*), it can be translated as follows:

(1') (x)(x must act such that [x does not kill innocent people])

The square brackets in rule (1') contain an ineliminable occurrence of the variable ‘ x ’ that is bound by the initial universal quantifier, and so (1') can be classified as agent-relative on these grounds. Thus, this translation avoids the first objection.

We can now look at the problem of finding an adequate translation of maximizing-rules. We saw above that translating (2) as either (2*) or (2[†]) is untenable. What we need is a better translation of (2) that adequately captures the notion of maximizing but, following McNaughton and Rawling’s key insight, does so with a verb that takes propositional content. I suggest the following:

(2') (x)(x must bring it about, to the greatest degree that x can, that [there is minimal killing of innocent people])

This rule adequately captures the notion of ‘maximizing’ because to maximize a quantity appears equivalent to bringing it about to the greatest degree you can that the world has the maximal amount of that quantity. Thus, consider the case of the Super Crime-Fighter discussed above. Clark Kent obeys rule (2) in this case, yet does not obey (2*). If (2') is an adequate translation of (2), then whenever Clark Kent obeys (2) he must also be obeying (2'). In Super Crime-Fighter, this is indeed the verdict that we get. When Clark Kent prevents as many killings of innocent people as he can, and yet fails to prevent every such killing, he is bringing it about, to the greatest degree that he can, that there is minimal killing of innocent people in the world.

Next, consider the case of unintentional maximizing, discussed above. An agent may maximize a quantity without intending or foreseeing that this is what her act will do. For example, Lex Luther might have a malicious intent to kill innocent people and yet, by accident, might act in a way that minimizes the killing of innocent people. In such a case, he minimizes the killing of innocent people yet fails to ‘ensure’ that he does this, because he actually intended to do otherwise. Thus, he obeys rule (2) without obeying (2*) or (2[†]). However, he does appear to obey (2') because he appears to (unintentionally) bring it about to the greatest degree that he can that there is minimal killing of innocent people. So, again (2') gives the same verdict as (2).

Reflection on further cases shows that (2) and (2') consistently give the same verdicts. Furthermore, (2') contains a proposition in square brackets that can be used to explain (2)'s status as agent-neutral. Thus, we appear to have found a way of translating maximizing-rules into semiformal notation such that they can correctly be classified by the presence or absence of the relevant kind of variable.

Given the translations that I have defended above for direct-rules and maximizing-rules, we can define agent-relative and agent-neutral rules as follows:

A rule expressing a moral requirement is agent-relative *iff*, when the rule is represented with the form ‘(x)(x must {verb} that [...])’, there is an ineliminable occurrence of ‘ x ’ in the square brackets that is bound by the initial universal quantifier.

A rule expressing a moral requirement is agent-neutral *iff* it is not agent-relative.

Applying these ideas to the other eight paradigm rules from [section 2](#), we get these:

- (3') (x)(x must act such that [(y)(y is x's child \rightarrow x cares for y)])
- (4') (x)(x must ensure that [(y)(z)(y is a child of z \rightarrow z cares for y)])
- (5') (x)(x must bring it about, to the greatest degree that x can, that [There is maximal well-being in x's life])
- (6') (x)(x must bring it about, to the greatest degree that x can, that [There is maximal well-being in the world])
- (7') (x)(x must bring it about, to the greatest degree that x can, that [(y)(y is x's friend \rightarrow there is maximal well-being in y's life)])
- (8') (x)(x must ensure that [(y)(y is not disloyal to the Supreme Leader)])
- (9') (x)(x must bring it about, to the greatest degree that x can, that [x does not violate rule (1)])
- (10') (x)(x must bring it about, to the greatest degree that x can, that [(y)(y does not violate rule (1))])

In these translations, all of the agent-relative rules contain an ineliminable occurrence of 'x' in the square brackets that is bound by the initial universal quantifier, whereas all of the agent-neutral rules do not. Furthermore, unlike rules (1*)–(6*) considered above, rules (1')–(10') are all accurate translations of the ordinary language rules (1)–(10). Thus, this revised account of the distinction avoids the two objections.


To conclude, let's consider what is attractive in the account of the agent-relative/agent-neutral distinction defended here. First, it correctly classifies all of the paradigm rules as either agent-relative or agent-neutral. McNaughton and Rawling's original account appeared to do this; however, it failed to adequately translate direct and maximizing rules.

Second, it helps us to make sense of Parfit's idea that rules can be classified as relative, or as neutral, by the ultimate aims that they give to agents. Perhaps it is not the case that all moral rules literally give aims to agents. However, we have seen that all rules can be translated such that agents must perform a type of act (such as ensuring, acting such that, bringing it about) with regard to some propositional content. Furthermore, we can see that, when the relevant variable occurs in a rule's content, this corresponds to that content being different for different agents, whereas the absence of the relevant variable corresponds to that content being the same for all agents. Therefore, if each rule's content is regarded metaphorically as what the rule is *aimed at*, then our formal account makes sense of the idea that some rules (but not others) give different aims to different agents.

Third, these ideas can be used to address Schroeder's scepticism about the distinction. Schroeder [2011] argues that if ought-statements are seen as a relation between agents and actions, rather than between agents and propositions, then all moral rules require the same thing of all agents, and thus the agent-relative/neutral distinction collapses. However, even if ought-statements are a relation between agents and actions, I have shown that the actions that they relate to agents may take propositional content and, following the basic strategy of McNaughton and Rawling, that formal features of that content can be used to identify a rule as agent-relative or as agent-neutral.

Given these various attractive features, I conclude that the account of the agent-relative/agent-neutral distinction that I have defended is the best available. As a formal account, it avoids some of the issues with Parfit's intuitive account (while also giving us a way of capturing the grain of truth in his account). It also avoids the problems that beset the other formal accounts in the literature—mainly the well-known objections to the accounts of Nagel, Sen, Pettit, and Skorupski, as well as the two objections that I develop against McNaughton and Rawling's account.¹³

ORCID

Matthew Hammerton  <http://orcid.org/0000-0003-4649-8206>

References

- Belnap, N., and M. Perloff 1988. Seeing to It That: A Canonical Form for Agentives, *Theoria* 54/3: 175–99.
- Broome, J. 1995. Skorupski on Agent-Neutrality, *Utilitas* 7/2: 315–17.
- Broome, J. 1999. Normative Requirements, *Ratio* 12/4: 398–419.
- Dougherty, T. 2013. Agent-Neutral Deontology, *Philosophical Studies* 163/2: 527–37.
- Dreier, J. 1993. Structures of Normative Theories, *The Monist* 76/1: 22–40.
- Hammerton, M. 2016. Patient-Relativity in Morality, *Ethics* 127/1: 6–26.
- Howard-Snyder, F. 1996. A New Argument for Consequentialism? A Reply to Sinnott-Armstrong, *Analysis* 56/2: 111–15.
- Hurka, T. 2003. Moore in the Middle, *Ethics* 113/3: 599–628.
- Louise, J. 2004. Relativity of Value and the Consequentialist Umbrella, *The Philosophical Quarterly* 54/217: 518–36.
- McNaughton, D. and P. Rawling 1991. Agent-Relativity and the Doing-Happening Distinction, *Philosophical Studies* 63/2: 167–85.
- McNaughton, D. and P. Rawling 1992. Honoring and Promoting Values, *Ethics* 102/4: 835–43.
- McNaughton, D. and P. Rawling 1993. Deontology and Agency, *The Monist* 76/1: 81–100.
- McNaughton, D. and P. Rawling 1995a. Value and Agent-Relative Reasons, *Utilitas* 7/1: 31–47.
- McNaughton, D. and P. Rawling 1995b. Agent-Relativity and Terminological Inexactitudes, *Utilitas* 7/2: 319–25.
- McNaughton, D. and P. Rawling 1998. On Defending Deontology, *Ratio* 11/1: 37–54.
- McNaughton, D. and P. Rawling 2002. Conditional and Conditioned Reasons, *Utilitas* 14/2: 240–8.
- Nagel, T. 1970. *The Possibility of Altruism*, Oxford: Clarendon Press.
- Nagel, T. 1986. *The View from Nowhere*, New York: Oxford University Press.
- Parfit, D. 1979. Prudence, Morality and the Prisoner's Dilemma, *Proceedings of the British Academy* 65: 539–64.
- Parfit, D. 1984. *Reasons and Persons*, Oxford: Clarendon Press.
- Pettit, P. 1987. Universality without Utilitarianism, *Mind* 96/381: 74–82.
- Pettit, P. 1997. The Consequentialist Perspective, in *Three Methods in Ethics: A Debate*, M.W. Baron, P. Pettit, M.A. Slote, Oxford: Blackwell: 92–174.
- Portmore, D.W. 2001. McNaughton and Rawling on the Agent-Relative/Agent-Neutral Distinction, *Utilitas* 13/3: 350–6.
- Portmore, D.W. 2013. Agent-Relative vs. Agent-Neutral, in *The International Encyclopedia of Ethics*, ed. H. LaFollette, Malden, MA: Wiley Blackwell: 162–71.

¹³ I am grateful to John Broome, Ryan Cox, Jamie Dreier, Daniel Nolan, Doug Portmore, and Jeremy Strasser for helpful comments on earlier versions of this paper. When I started doing work related to the agent-relative/agent-neutral distinction, some of my more formally oriented colleagues expressed the concern that there is no clear distinction here at all. Their scepticism led me to the present project, and thus I thank them for impressing on me the need for a better account of the distinction.

- Ridge, M. 2005. Reasons for Action: Agent-Neutral vs. Agent-Relative, *The Stanford Encyclopedia of Philosophy* (Winter 2005 Edition) ed. E.N. Zalta, URL = <http://plato.stanford.edu/archives/win2005/entries/reasons-agent>
- Schroeder, M. 2011. Ought, Agents, and Actions, *The Philosophical Review* 120/1: 1–41.
- Sen, A. 1982. Rights and Agency, *Philosophy and Public Affairs* 11/1: 3–39.
- Skorupski, J. 1995. Agent-Neutrality, Consequentialism, Utilitarianism: A Terminological Note, *Utilitas* 7/1: 49–54.
- Skorupski, J. 1996. Neutral Versus Relative: A Reply to Broome, and McNaughton and Rawling, *Utilitas* 8/2: 235–48.
- Wedgwood, R. 2006. The Meaning of 'Ought', *Oxford Studies in Metaethics, Vol. 1*, ed. R. Shafer-Landau, Oxford: Oxford University Press: 127–60.