

RADICAL EVALUATIVE IGNORANCE

Martin Peterson

Texas A&M University

1 Introduction

If you can't figure out if Kandinsky's *Squares with Concentric Circles* is beautiful or ugly, or if you don't know if euthanasia could ever be morally permissible, then your ignorance is evaluative rather than factual. Evaluative ignorance is either *internal*, *external*, or *radical*. Internal evaluative ignorance arises if you are ignorant of your own evaluative attitudes, such as your preferences or desires. It is widely agreed that we can be ignorant of our future and past evaluative attitudes, but not everyone believes we can be ignorant of our present ones.¹ I shall not discuss internal evaluative ignorance here.

External evaluative ignorance is similar to its internal counterpart except that the evaluation does not *merely* concern the agent's subjective attitude.² Consider, for instance, moral uncertainty. Whether it would be morally permissible for a professor to lie about what the class should read in preparation for the final exam ("You must read the whole book!") depends on what moral theory happens to be correct. As long as not all moral theories entail the same verdict about what the professor should do, and given that the professor is ignorant of which moral theory is correct, then this is an example of external evaluative ignorance; this is because the correctness of a moral theory does not depend, at least not always and entirely, on the agent's subjective evaluative attitude.³

The topic of the present chapter is the third type of evaluative ignorance, which I call *radical* evaluative ignorance. By radical evaluative ignorance I mean ignorance about what *source of normativity* is, or is not, applicable to some issue. Some normative verdicts are determined by morality, but others are determined by aesthetical, epistemic or self-interested considerations. If

¹ See Peterson (2005).

² For reasons that will become clear shortly, it is helpful to stipulate that external evaluative ignorance refers exclusively to cases in which the agent's ignorance is not radical.

³ See, for instance, Lockhart (2000), Sepielli (2013), and Gustafsson and Torpman (2014).

you are ignorant of what source of normativity is applicable in some situation, then your evaluative ignorance is radical. Consider the following example: Alice has one thousand dollars left at the end of the month to spend on whatever she pleases. She is keen on buying a new computer, although there is nothing wrong with her old one. From a self-interested point of view, this is what she ought to do. However, Alice is also convinced that the morally right thing to do would be to donate her one thousand dollars to charity. There are many poor people who would need the money more than Alice. Just like so many others of us, Alice wonders what she ought to do *all things considered*. That is, Alice is ignorant of what *type* (or types) of normative reasoning is applicable to her problem, and this is what makes her ignorance radical. Is this a situation in which only moral considerations are relevant, or are other types of normative reasons also applicable? What should Alice do, all things considered, if she is a normatively conscientious person who gives due attention to all her normative obligations?

In what follows I will leave it open whether evaluative ignorance (of the internal, external, or radical type) is best characterized in noncognitivist terms, or as the absence of knowledge of some evaluative proposition *e*, or the absence of a true belief that *e*.⁴ The traditional metaethical debate over cognitivism and noncognitivism will not be discussed here, nor shall I try to offer any general account of what ignorance is, and is not. The question of how a normatively conscientious agent should deal with radical evaluative ignorance arises no matter what minimally plausible view one takes on the nature of ignorance and the debate over cognitivism and noncognitivism.

The aim of this chapter is to defend two claims. My first claim is that we are sometimes confronted with radical evaluative ignorance, meaning that there *are* cases in which we are ignorant of what source of normativity is applicable to an issue. My argument for this claim is, essentially, that it would be a mistake to think that there is only one source of normativity, or that one source overrides all the others. If two or more sources have to be considered by normatively conscientious agents, it would be overly optimistic to think that no such agent would *never* be ignorant of which source determines our normative verdicts. My second claim, which is more complex, is a positive proposal for how to deal with radical evaluative ignorance. To put it briefly, I propose that a normatively conscientious agent confronted with radical evaluative ignorance

⁴ For an interesting discussion of whether ignorance is the lack of knowledge of something or the absence of a true belief, see the debate between Rik Peels and Pierre Le Morvan; e.g. Peels (2010) and Morvan (2012).

should compare how *similar* her case is to prototypical examples of each source of normativity, and then apply the type of normative considerations that are applicable to the prototypical case that is most similar to the one she is facing.

2 Is Radical Evaluative Ignorance Possible?

For present purposes, an entity *S* is a *source of normativity* if and only if there is at least one possible world *w* and one agent *A* such that *S* makes a difference to what *A* in *w* ought to do (including what to say or believe) all things considered. In this definition the phrase “making a difference” means that the all-things-considered ought would have been different on at least one occasion for *A* in *w* if *S* would have been silent or inapplicable in that situation. Imagine, for instance, that the morally right thing for you to do right now in the actual world is to donate most of your money to Oxfam. Then, if this moral verdict influences what you ought to do all-things-considered right now in the actual world, it follows that morality is a source of normativity.

If there exists only one source of normativity, and we know it, then it seems that we cannot be ignorant of what source Alice should consider in her deliberations.⁵ The one and only source would, of course, be applicable to Alice’s as well as every other case. So to show that radical evaluative ignorance is possible, I first have to demonstrate that we have good reason to think that different sources of normativity apply to different situations.

Scholars who believe that there exists only one source of normativity defend what I shall call a *unary* account of normativity. Consider, for instance, Torbjörn Tännsjö’s claim that, “there exists exactly one source of normativity. And this is the one I speak of as the *moral* one.”⁶ Other philosophers have proposed alternative ways of thinking about the relation between different sources of normativity. Some argue that one source (morality) always override all other sources (such as self-interest).⁷ Other agree with Tännsjö that there is only one source, or that all the different sources always coincide, meaning that there is no genuine conflict between them. It has

⁵ The same applies if there exists more than one source, but we know that one source always overrides all other sources. The argument outlined in this section applies *mutatis mutandis* to this type of view as well.

⁶ Tännsjö (2010: 51).

⁷ See, for instance, Singer (2004).

also been suggested that although there exist a plurality of sources, none of them overrides the others.⁸

That one source of normativity *overrides* another means that it is normatively superior in the sense that the overriding source trumps the other source. Another way of putting this is to say that the all-things-considered ought is determined, on all occasions and in all possible worlds, by the overriding source instead of the other source. Authors who believe that morality and self-interest are incomparable conclude that because neither source trumps the other, no all-things-considered normative verdict can ever be reached. On this view, something is merely right, good, better or permissible *with respect to* some source of normativity, but not right, good, better or permissible *simpliciter*.

I shall now introduce a technical term, which will be frequently mentioned in the following sections. By stipulation:

Rigidity

A source of normativity *x* is *rigid* if and only it holds in all cases, in all possible worlds, for some other source *y* that (i) *x* overrides *y*, or (ii) *y* overrides *x*, or (iii) *x* and *y* always yield identical verdicts, or (iv) *x* and *y* are incomparable.

The reason for using the term *rigid* for describing this type of structural view is that, on any such view, the relation between *x* and *y* is always the same. No matter how strong or weak *x* and *y* are, the all-things-considered normative verdict will remain unaffected.

Note that if *x* and *y* are rigid sources of normativity, and we know which of the four conditions it is that makes this the case, then there is little room for radical evaluative ignorance, as indicated earlier.⁹ A straightforward way to show that radical evaluative ignorance is more than a mere conceptual possibility is, therefore, to show that there exist at least two non-rigid sources of normativity.

⁸ This view is defended by Sidgwick (1874) and Copp (1997).

⁹ Rik Peels has pointed out to me that it is conceptually possible that we know that the disjunction in the definition of rigidity is true, but we do not know *which* disjunction it is that makes the disjunction true. Although I agree that this is a conceptual possibility, I have little advice to offer to agents facing this somewhat odd type of radical evaluative ignorance.

To keep things simple, I will in what follows discuss just two putative sources of normativity: morality and self-interest. The points I make about morality and self-interest also apply to epistemic and aesthetic normativity. In *On What Matters*, Parfit discusses Sidgwick's hypothesis that morality and self-interest are incomparable. Let us call Parfit's interpretation of this view *Sidgfit's dualism*. Parfit formulates Sidgfit's dualism as a claim about reasons. Sidgwick himself used a different terminology, but little or nothing is lost by adopting Parfit's vocabulary:¹⁰

Sidgfit's Dualism

Impartial and self-interested reasons are *wholly* incomparable. *No* impartial reason could be either stronger or weaker than *any* self-interested reason.¹¹

Advocates of Sidgfit's dualism believe that self-interest and morality (that is, impartial reasons) are rigid sources of normativity. This is because the last disjunct of the definition of rigidity stated above is satisfied.

Ethical egoists also defend a rigid position. They believe that morality and self-interest always coincide: The morally right option is to do what is best from the agent's self-interested point of view. Aristotle is sometimes interpreted as an advocate of a form of ethical egoism. His argument for the view that morality and self-interest always coincide is that "happiness is an activity of the soul in accordance with perfect virtue".¹² This quote seems to entail that it is in your self-interest to be happy, and you can only become happy by becoming virtuous.

Having said that, the most straightforward examples of rigid theories are, of course, theories in which one source of normativity is claimed to override all other sources. Unlike Sidgwick, contemporary utilitarians routinely stress that morality overrides self-interest.¹³ If you face a choice between either doing something that would make yourself a million units better off without affecting anyone else, or increasing the sum total of utility in the world by one million and one units by doing something that would not make you any better off, you ought all-things-considered perform the latter option.

¹⁰ In his (1874), Sidgwick does not use the terms "impartial reason" and "self-interested reason" at all. The term "a reason" is used only nine times, mostly in other contexts.

¹¹ Parfit (2011: 32).

¹² Aristotle, NE:I.

¹³ See, for instance, Singer (2004).

Somewhat surprisingly, Nietzsche's theory of normativity is also rigid in this sense. However, contrary to contemporary utilitarians, he maintains that self-interest overrides morality under all circumstances:

An "altruistic" morality -- a morality in which self-interest withers away -- remains a bad sign *under all circumstances*... The best is lacking when self-interest begins to be lacking.¹⁴

It is beyond the scope of this paper to discuss what led Nietzsche and the other authors quoted above to hold the rigid views they do. All I take these examples to show is that several influential thinkers have in fact proposed and defended rigid theories of normativity. In what follows I will first argue that all unary and rigid theories face a serious challenge, which gives us reason to believe that radical evaluative ignorance is more than a mere conceptual possibility.

3. The No Difference Argument

Unary as well as rigid theories of normativity lead to absurd conclusions if we vary the *strength* of reasons produced by different types of sources. Let me first explain why this is a problem for rigid theories, before I go on to show that a similar objection can also be raised against unary theories.

Let us suppose, for the sake of the argument, that x is a rigid source of normativity with respect to y . Then imagine a situation in which you have a strong reason to do what is required by x , but a weak reason to do something else required by y . For the sake of the argument, we stipulate that the difference in strength between the two reasons be huge.¹⁵ Now compare this case to another, in which you have a strong reason to do what is required by y but a weak reason to do what is required by x . The difference in strength between the two types of reasons is, again, huge. By definition, every rigid theory of normativity entails that these huge differences make *no*

¹⁴ Nietzsche (1889: 35), my italics.

¹⁵ Note that the term "difference in strength" does not commit us to comparisons across different sources of normativity. If you have a weak moral reason to do something but a strong self-interested reason to not do it, then your moral reason is weak *compared to other moral reasons*, while your self-interested reason is strong *compared to other self-interested reasons*.

difference to the all-things-considered ought. However, according to what we may call the No Difference Argument this is deeply counterintuitive.

Parfit discusses a particular version of the No Difference Argument in his analysis of Sidgwick's dualism:

Suppose we are choosing between some architectural plans for some new building... If economic and aesthetic reasons [i.e. sources of normativity] were wholly incomparable, it would therefore be true both that

(1) we could rationally choose one of two plans because it would make this building cost one dollar less, even though this building would be very much uglier,

and that

(2) we could also rationally choose one of two other plans because it would make this building slightly less ugly, even though this building would cost a billion dollars more.¹⁶

According to Parfit, it is not true that a single dollar is *not* outweighed by a huge difference in ugliness in the first case, and that a small difference in ugliness is *not* outweighed by a billion dollars in the second case. Therefore, the two sources of normativity are *not* wholly incomparable.

Advocates of the No Difference Argument believe that Parfit's argument can be generalized to other sources of normativity. If a difference is large enough it should matter, and this does not just hold true for economic and aesthetic considerations. This is, on the contrary, a general insight that holds for all sources of normativity. Consider the following example and imagine for the sake of the argument that morality and self-interest are rigid sources of normativity:

(1') you could rationally choose to save yourself from dying in a fire, even though you could have stayed in the burning house and instead have taken some action that would have killed

¹⁶ Parfit (2011: 132-3).

yourself but saved a stranger, who would have experienced a millionth of a unit of more wellbeing in the future.

Also imagine that:

(2') you could rationally choose to treat yourself with one additional drop of wine in the bar tonight, which you have a weak self-interested reason to do, even though you could instead have saved a billion people from starving to death by pressing a green button.

In case (1') the phrase "a millionth of a unit of more wellbeing" is a placeholder for any feature that makes saving a stranger just a tiny bit better from a moral point of view according to one's favorite moral theory. The consequentialist flavor of (1') reflects my own (MP's) moral outlook but is not essential to the argument.

Now, starting from what seems to be uncontroversial intuitions about (2') and (1'), we can construct the following trilemma: If morality overrides self-interest we must *reject* (1'), which is counterintuitive, and if self-interest overrides morality, we must *accept* (2'), which is also counterintuitive. Moreover, if morality and self-interest are incomparable, and given that it is rational to choose any of two incomparable alternatives, we must accept *both* (1') and (2'). However, we have already pointed out that it is counterintuitive to accept (2'), so therefore this option is no more attractive than claiming that self-interest overrides morality. Finally, if we reject the premise of the third horn of the trilemma, according to which it is rational to choose any of two incomparable alternatives, we run into an analogous problem. If we instead claim that it is not rational to choose *any* of two incomparable alternatives, we must *reject* both (1') and (2'). However, we have already pointed out that it is counterintuitive to reject (1'), so this maneuver is no more attractive than claiming that morality overrides self-interest.

Note that this trilemma is perfectly general. We can always modify (1') and (2') such that no matter which particular theories of self-interest and morality we happen to believe in, it holds true that in (1') we have a *strong self-interested reason* to do what is required by our theory of self-interest, but a *weak moral reason* to do something else required by our moral theory. We then compare the first case to (2'), in which we have a *strong moral reason* to do what our moral theory requires of us, but a *weak self-interested reason* to do something else.

All that said, the scope of the trilemma proposed here is restricted to rigid views. Authors who defend unary views would of course object that self-interest and morality cannot come apart in the way stipulated in the premises of the trilemma. It takes little effort to see that if there exists only one source of normativity, then, for any sources x and y , it will always be the case that x and y are identical.

What could advocates of the No Difference Argument say in response to the claim that there exists only one source of normativity? Note that to refute this view it suffices to find a *single case* in which self-interest and morality come apart. Consider the following counterexample:

The Green Button Case

You will die tomorrow, and you know this. Your death is inevitable, and it will, unfortunately, be very painful unless your doctor gives you some morphine. On the desk in front of you is a green button. If you press the button, all wars will come to an end, world poverty will be eradicated, and all diseased people in the world (except you) will be cured. However, if you refrain from pressing the green button no wars will stop, world poverty will continue to haunt the world, and no one will be cured of any disease. The only upside of not pressing the button is that your doctor will then give a dose of morphine, which will make your inevitable death less painful.

In the Green Button Case, morality and self-interest do seem to come apart. You have a strong moral reason to press the green button because that would make the world as a whole much better. At the same time, you have at least *some* self-interested reason to refrain from pressing the green button. Therefore, if advocates of the unary view are right that there is only one source of normativity, we face the following dilemma: We must either insist that (i) your moral and self-interested reasons coincide in the Green Button Case, which seems implausible, or we must claim that (ii) you either have no moral reason to push the button *or* no self-interested reason to refrain from pushing the button. All these conclusions are deeply counterintuitive.

The Green Button Case is designed to refute unary theories of the type proposed by Tännsjö and mentioned in Section 2.¹⁷ Note, however, that nothing hinges on the assumption that what appears to be different sources of normativity are in fact one and the same. The Green Button Case

¹⁷ Tännsjö, *Ibid.*

works equally well if we accept the somewhat less extreme view that morality and self-interest are separate sources of normativity that happen to entail the same normative verdicts about all cases; the objection raised against Tännsjö's theory will also arise under this alternative assumption. The reason for this is that no matter which minimally plausible moral theory you believe in it seems clear that you ought to press the green button, and no matter which minimally plausible theory of self-interest you believe in it seems clear that you ought not to press the button.

Strictly speaking, the Green Button Case is an independent supplement to the No Difference Argument. The Green Button Case is merely designed to refute unary theories, which are not discussed by Parfit. However, the gist of the Green Button Case is the same as that of the No Difference Argument. Any minimally plausible view about normativity should be sensitive to the relative strength of the normative reasons that speak for and against the alternatives open to us. The No Difference Argument and the Green Button Case together bring out the implausible consequences that follow if we deny this.

4 Mainstream Decision Theory Cannot Solve the Problem

If the argument of the preceding section is correct, then radical evaluative ignorance is more than a mere conceptual possibility. It is likely that we actually face this type of ignorance from time to time, because we have reason to believe that there exist two or more non-rigid sources of normativity. Alice's decision to either spend the one thousand dollars she has left at the end of the month on a new computer or donate the money to famine relief is as a good example of this.

Could we handle radical evaluative ignorance by applying any of the standard techniques of decision theory? Consider Figure 1, which summarizes the situation faced by Alice. She can either buy a new computer or donate \$1000 to a charity organization. It is helpful to keep things simple by supposing that the only relevant states of the world are the ones in which either morality or self-interest is the relevant source of normativity. Let us also suppose that Alice is able to represent her subjective degree of belief in each of the two states in a manner that satisfies the axioms of the probability calculus (which means that, in a strict sense, Alice is not facing a decision under ignorance). However, what makes it difficult to apply the standard ideas from decision theory to Alice's decision is that there seems to be no meaningful way in which she could rank the possible outcomes. In order to calculate the expected "value" of each act she would have to measure how good or bad the four possible outcomes are relative to each other.

	Morality is the relevant source of normativity (pr=0.7)	Self-interest is the relevant source of normativity (pr=0.3)
Buy new computer	<i>Morally wrong act</i>	<i>The act is rational</i>
Donate \$1000	<i>Morally right act</i>	<i>The act is irrational</i>

Figure 1. An example of radical evaluative ignorance

To put it briefly, the problem is that it seems impossible even in principle to make "inter-source" comparisons of normativity. How good or bad would it be to perform the morally right act (that is, donate \$1000 given that morality is the relevant source of normativity) compared to acting in accordance with her self-interest (buy a new computer given that self-interest is the relevant source of normativity)? The problem is not that it is hard to know this, or that we have not yet been able to carry out the required measurements. The problem is that the very comparison seems to have no meaning.

Similar doubts about the possibility of inter-theoretical comparisons of moral value have been raised in the literature on moral uncertainty.¹⁸ The background to that discussion is, in turn, the debate over whether it is possible to make interpersonal comparisons of utility. Critics of utilitarianism argue that because the utilitarian theory presupposes that interpersonal comparisons are possible, and such comparisons are impossible, the utilitarian criterion of moral rightness has no meaning.

In the literature on moral uncertainty it has been proposed that the best response to the problem of inter-theoretical comparisons is to replace the principle of maximizing expected value by the principle holding that it is morally conscientious to act in accordance with the theory one has most credence in. Here is an example: If your credence in the utilitarian theory is higher than your credence in Kantianism, then you should act as if you were *entirely sure* that the utilitarian theory is correct. Needless to say, we could apply a similar principle to radical evaluative ignorance. The normatively conscientious choice for Alice would then be to assume that morality is the relevant source of normativity and that the normatively conscientious option would,

¹⁸ This point is made by Lockhart (2000), which is the point of departure for many recent discussions on moral uncertainty. See also Gustafsson & Torpman (2014).

therefore, be to donate \$1000 to charity. This principle does not require any intertheoretical comparisons.

However, a problem with this proposal is that Alice's decision then becomes sensitive to the *individuation* of sources of normativity. For instance, for all Alice knows the moral source can be split up into two "sub-sources": a utilitarian source and a Kantian one. What it would be normatively conscientious for Alice to do would then depend on how she individuates all these sources (and recall that both the utilitarian and Kantian theory comes in many different versions). A possible response to this objection could be to argue that some principle for individuating sources is better than all others. Here is a possible suggestion: "Treat two sources as different if and only if they sometimes yield different verdicts". I leave it to the reader to adjudicate whether this is a reasonable individuation principle. However, an additional problem with replacing the principle of maximizing expected value by the principle holding that it is normatively conscientious to act in accordance with the theory one has most credence in, is that the agent will then sometimes act *as if* a source she believes to a very low degree to be applicable is, in fact, the one and only applicable source. In principle, the agent may end up acting *as if* she fully believed that a source were the relevant one even though her credence in that hypothesis was very low. If Alice's credence is, say, 0.0001 that the moral source applies to her case, then it is not normatively conscientious for her to *act as if* her credence in this hypothesis is 1.

Rik Peels has pointed out to me that an additional problem with the idea that it is conscientious to act in accordance with the theory one has most credence in, is that this does not reduce one's ignorance. Perhaps the agent will end up acting in accordance with only one source of normativity, but she has not eliminated any evaluative ignorance.

5 Summary of My Proposal

My proposal for how to deal with radical evaluative ignorance draws on an influential line of research in cognitive science initiated by Eleanor Rosch (1973, 1974) and developed further by Peter Gärdenfors (2000, 2014). Before I explain how this work is helpful for understanding abstract philosophical issues about normativity, I would like to summarize Rosch's and Gärdenfors' work.

Rosch criticizes the Aristotelian hypothesis that concepts are demarcated by some set of necessary and sufficient conditions that are fulfilled by all items that fall under a concept.

According to Aristotle, a penguin counts as a bird if and only if it fulfills all the necessary and sufficient conditions for being a bird. In her work, Rosch challenges the descriptive accuracy of Aristotle's theory and claims that this is a poor account of how we *actually* categorize different items. According to Rosch, human cognizers rather count a penguin as a bird because it is more similar to a *prototype* bird than to any other prototype for any other animal, such as a shark or polar bear.

Gärdenfors (2000, 2014) develops Rosch's work further and proposes a theory of *conceptual spaces* in which the degree of similarity between a prototypical bird and a penguin is represented as the distance in some (Euclidean or non-Euclidean) geometric space. Gärdenfors points out that it is much easier for a cognizing agent to compare new items with some nearby prototype instead of categorizing them by applying some set of necessary and sufficient conditions as proposed by Aristotle. Consider, for instance, a child who is attempting to learn whether a penguin is a bird or not. If Aristotle is right, the child would first have to learn a large number of necessary and sufficient conditions for different animals and then correctly apply the conditions for being a bird to the penguin. The problem with this Aristotelian theory is that it makes it very difficult to learn new concepts. Young children learn thousands of new concepts in a very short period and it is not plausible to believe that they would have the cognitive capacity to store information about the corresponding number of necessary and sufficient conditions in their brains. In Gärdenfors' theory of conceptual spaces much less information is needed for learning new concepts. All a child (or adult) has to be able to do is to compare *how similar* a new item is to the prototypes for the concepts she already knows. If the new item (a penguin) is similar enough to a prototype (a prototype for a bird), then the new item falls under that concept. So instead of storing information about necessary and sufficient conditions, the child just has to store information about a small number of prototypes and the distances between various items and the prototypes. This cognitive economy may explain why evolutionary processes may have favored organisms who operate in the manner proposed by Rosch and Gärdenfors instead of the much more complicated way suggested by Aristotle.

Rosch and Gärdenfors do not discuss how we learn philosophical concepts.¹⁹ Their aim is to describe and explain how people develop ordinary empirical concepts such as "tree", "house"

¹⁹ Paul Churchland has discussed the implication of prototype theory for virtue ethics; see Churchland (2007, Ch. 3).

and “bird”. Needless to say, the aim of the present paper is not to explain how we learn such concepts, nor how beliefs about normativity are *actually formed* in our minds. What is at stake here is a claim about how our thinking about normativity *could* and *should* be shaped when we face radical evaluative ignorance, not any claim about how we *actually* form these views. It is thus irrelevant whether the cognitive theory developed by Rosch and Gärdenfors is descriptively accurate.

In analogy with the prototype theory in cognitive science, I propose that an agent confronted with radical evaluative ignorance should compare the situation she is facing to situations that are prototypical for each and every source of normativity. By a prototype I mean a case in which it is paradigmatically clear that the source in question determines the all-things-considered ought, unaided by other sources. Imagine, for instance, that you can save ten million people from starving to death at minimal cost to yourself. This could be a candidate for a prototype for morality; and a prototype in which self-interest overrides morality could be the case in which you are offered to choose between a window or aisle seat when checking in for your next flight. In the latter case it is, under normal circumstances, paradigmatically clear that you are free to let your preference be decisive without considering the interests of others.

How can the agent facing radical evaluative ignorance make use of these prototype cases for figuring out what to do? According to the view I propose, the agent should compare how *similar* her case is to the prototypes for each source of normativity. To be more precise, I propose that it is normatively conscientious for the agent to assume that the all-things-considered ought is determined by its degree of similarity to the most similar prototype. If, for instance, Alice’s case is more similar to a prototype in which morality overrides all other sources, then it is normatively conscientious of Alice to do whatever morality demands of her. It is thus the degree of similarity to nearby prototypes that determines what the normatively conscientious agent has most reason to do.

The theory I propose can be illustrated in a Voronoi diagram.²⁰ A Voronoi diagram divides space into a number of regions such that each region consists of all points that are closer to a

²⁰ Note that nothing hinges on the printed illustrations. A Voronoi diagram is an abstract geometric object, just like a perfect circle. The geometric account could, at least in principle, be stated without using any illustrations at all.

predetermined seed point (prototype) than to any other seed point. Figure 1 shows a Voronoi diagram with five seed points. Within each region belonging to a given seed point (source of normativity), the normatively conscientious conclusion is determined by the source of normativity corresponding to the seed point in question. Hence, in all cases that are more similar to, and hence closer to, a prototype for morality than to any prototype for some other source of normativity, morality overrides all other sources of normativity.

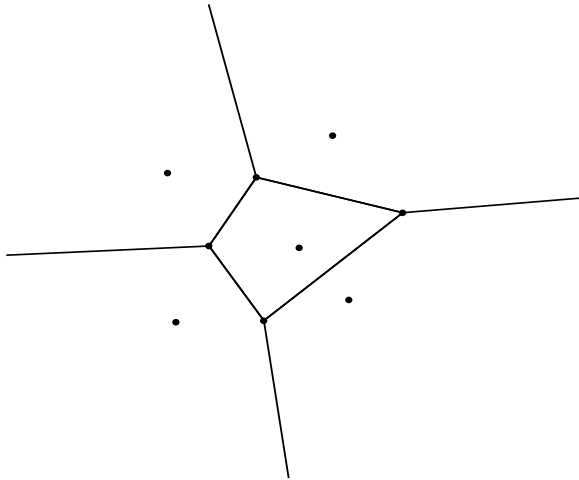


Figure 2. A Voronoi diagram with five seed-points.

In Figure 2, similarity is represented by the Euclidean distance between two cases. The more similar they are, the closer are the points in the figure. A possible objection to this representation is that the Euclidean measure is just one of many alternative ways of representing similarity. As frequently noted by philosophers of science, the Euclidean distance measure is not the only measure on the market. So why should we prefer a cardinal measure of similarity over ordinal measures? Moreover, why is the Euclidean measure the most appropriate cardinal measure? A third and final question is how we should identify the relevant dimensions of the two (or 3, 4, ... n) dimensional Euclidean space.²¹

Let me discuss each question in turn. Briefly put, the answer to the first question is that we should prefer cardinal measures over ordinal ones whenever the information required for

²¹ For useful overviews of the literature on these and many related issues, see Gärdenfors (2004) and (2014). My intellectual debt to Gärdenfors should be obvious by now for readers who are familiar with his work.

constructing such a measure is available. Cardinal measures are preferable over ordinal ones simply because they permit us to make more nuanced comparisons. That said, it is, of course, true that cardinal comparisons are not always possible. Some normatively relevant features of some cases cannot be measured on any cardinal scale, meaning that no cardinal measure could give an accurate representation of how similar such cases are.

Let me now turn to the second question: When and why is the Euclidean measure the most appropriate cardinal measure? The answer is that the choice between different measures of similarity depends on the nature the cases under consideration. As frequently pointed out in the literature, there is a large number of alternative measures to choose from.²² It would be naïve to claim that the Euclidean distance measure should always be adopted. Sometimes other measures might be more appropriate, depending on the nature of the properties believed to be normatively relevant in the cases we wish to compare.

The upshot is that there is no universal measure of similarity that can or should be applied to each and every possible case. However, in what follows, I will adopt the Euclidean distance measure unless otherwise stated. This is not because this measure applies to each and every conceivable case, but rather because the Euclidean measure is easy to understand and illustrate. It is therefore reasonable to base most of the examples in the following sections on the Euclidean measure, although it is, of course, worth keeping in mind that none of the core ideas of the non-rigid theory depend on which distance measure is chosen. All that is required is that we can make sense of the idea that some cases are more similar to each other than to any other case.

Let me finally discuss the first and final question asked above: How do we determine the relevant dimensions of the Voronoi diagram? A naïve but sometimes useful strategy is to identify the relevant dimensions by studying the information available to the agent and then decide *ex ante* what the relevant dimensions are. If the naïve strategy is pursued, we *first* identify the relevant dimensions and *thereafter* compare how similar the cases are along these dimensions.

There is also a more sophisticated strategy for identifying the relevant dimensions called multidimensional scaling (MDS). This is a standard technique used by psychologists and others for representing similarities across a set of data points. An important feature of MDS is that dimensions are identified *after* data has been collected. In the present context, this means that we start with a set of judgments about cardinal distances (which may be neither Euclidean nor metric)

²² For an overview, see Tverksy and Gati (1978).

between each pair of cases under consideration. We then construct the dimensions such that the fit between the degree of similarity between the data points and the formal representation is as tight as possible. This is a well-known optimization problem for which several computer algorithms are available.

In MDS, it is up to the person doing the analysis to decide how *many* dimensions should be considered in the optimization process. In MDS the dimensions have no meaning that is independent of the data points, they merely reflect the relative positions of the data points fed into the algorithm. This means that it is up to the researcher to propose a plausible interpretation of the dimensions. The larger the number of dimensions is, the more accurate the fit will be. However, if the number of dimensions is very large, it becomes harder to propose meaningful interpretations of the dimensions.²³

6 Conflicting Sources of Normativity

So far I have assumed that the conclusion of the normatively conscientious agent is entirely determined by a *single* source of normativity, as illustrated in Figure 2. The only exceptions are cases located exactly on the border between two regions. Such cases are equally far away from two prototypes, so they are strictly speaking influenced by two sources of normativity. This is, however, a quite implausible limitation of my view. A more plausible version should, arguably, be able to make sense of the idea that sometimes the normatively conscientious agent should conclude that *many* sources contribute to the all-things-considered ought.

Consider, for instance, the choice faced by Alice in the introduction. This might very well be a case in which *both* morality and self-interest contribute to the all-things-considered ought, although there is no reason to think that the case is located *exactly* on the border between two regions. To see why, suppose that we alter the amount of money Alice has left at the end of the month from \$1000 to \$999 and that this is sufficient for buying a new computer. If the initial version of the example is exactly on the border, then the new version cannot also be exactly on the border, because now the strength of the self-interest reason is the same while the moral reason is

²³ It is worth mentioning that I have done some experimental work in which I have applied MDS to (moral) similarity comparisons to data obtained from 240 academic philosophers. In this so far unpublished study it was possible to obtain a meaningful two-dimensional interpretation. Although inconclusive, this gives us some reason to believe that the conceptual space approach works also in normative contexts.

somewhat weaker. However, in both versions it might be plausible to maintain that both morality and self-interest contribute to the all-things-considered ought.

The best way to accommodate the thought that more than one source of normativity may contribute to the all-things-considered ought is to deny that each source has *exactly* one prototype. If we instead believe that some sources have *several* prototypes, then some regions in the Voronoi diagram may overlap each other. Figure 3 illustrates an example with five sources of normativity, in which three of the sources have two prototypes.

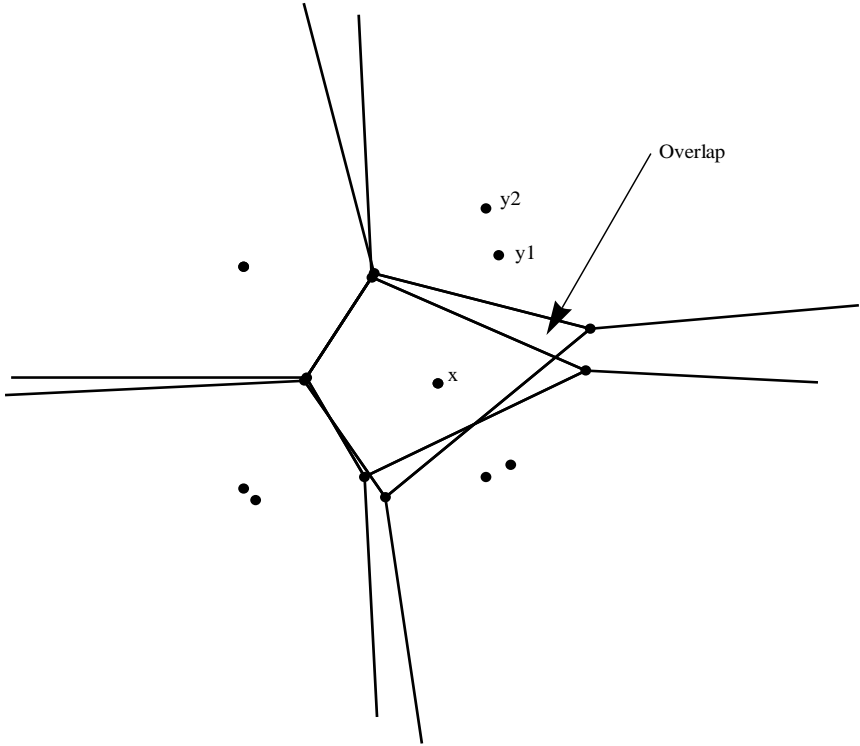


Figure 3. Overlapping regions of normativity.

Each non-prototypical case in the region marked by an arrow in Figure 3 is closest to prototype y_1 when compared to prototype x but closer to x when compared to the second prototype y_2 . The set of non-prototypical cases for which this holds true defines one of several overlapping regions in Figure 3. In all such regions, more than one source of normativity contributes to the all-things-considered ought. Before I discuss what the normatively conscientious agent should do when more than one source of normativity contributes to the all-things-considered-ought, it is helpful first to

consider the conditions under which this type of case can arise. Consider Figure 4, which depicts a Voronoi diagram with only two sources of normativity. The rightmost region has two prototypes, y_1 and y_2 , and the dashed line denotes the Voronoi border between y_2 and x . The corresponding Voronoi border between y_2 and x is depicted by a solid line in the diagram.

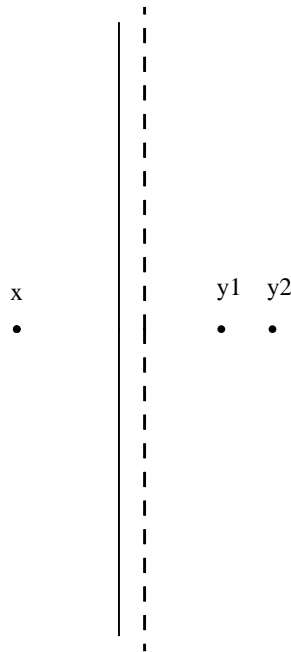


Figure 4. Conflicting sources of normativity.

More than one source of normativity contributes to the all-things-considered ought in all the cases that lie between the solid and dashed lines. This is because the non-prototypical cases between the dashed and solid lines are closer to the left-most region when x is compared to y_2 , but closer to the right-most region when y_2 is compared to x . In what follows the term “normative grey area” will be used for referring to cases in which more than one source of normativity contributes to the all-things-considered ought, as well as to some similar types of cases. To be more precise, a normative grey exists whenever the following conditions are met: (i) There exist at least two sources of normativity. (ii) At least one source of normativity has more than one prototype. (iii) A source of normativity x contributes to the all-things-considered ought in a non-prototypical case c if and only if, for some prototype p_x for x , it holds that $|c - p_x| < |c - p_y|$ for every other source y and prototype p_y .

Conditions (i) and (ii) are straightforward, but condition (iii) is somewhat less transparent. Briefly put, it states that a source of normativity x contributes to the all-things-considered ought in some non-prototypical case c if and only if c is more similar to at least one of the prototypes for x , compared to *some* prototype for every other source of normativity.

A possible way to escape the conclusion that more than one source of normativity contributes to the all-things-considered ought, even when some sources have more than one prototype, is to argue that when two or more prototypes for the same source is available, it is only the one that is closest to the non-prototypical case under consideration that contributes to the all-things-considered ought. In the example illustrated in Figure 4, this would entail that source y_2 would not contribute to the all-things-considered ought of any non-prototypical cases, and the only cases in which more than one source would contribute to the all-things-considered ought would be the cases in which the non-prototypical cases lie exactly on the border between x and y_1 . Although this would no doubt yield a less complex picture of the normative landscape, it seems that such a position fails to do justice to the idea that all prototypes for each source matter. Because, for instance, y_1 and y_2 are prototypes for the same source, it seems that each of them should contribute to the all-things-considered ought. If some prototype for some source of normativity turns out to sometimes be totally inert, because some other prototype for the very same source of normativity happens to be located closer to the non-prototypical case under consideration, it seems that the former prototype was after all not a prototype. To be a prototype for a source of normativity is to be a case that *defines* the source of normativity in question. No such prototype for a source of normativity can be ignored; therefore, if one prototype for a source of normativity contributes to the all-things-considered ought in some case, then so do all prototypes for that source.

That said, it is plausible to think that the distance between a non-prototypical case and the nearby prototypes affect *how much* influence each prototype should be allowed to have on the all-things-considered verdict. Cognitive scientists have explored the hypothesis that the influence of a prototype decreases as the distance to a non-prototypical case increases. This relationship may be non-linear. Hampton proposes an S-shaped function:²⁴ (i) For objects that are close to the prototype the influence does not decrease very much at the beginning, and (ii) halfway between the prototype and its Voronoi border the influence is decreasing rapidly, and (iii) close to the Voronoi border the influence is very low and slowly decreases to zero exactly at the border. In

²⁴ See Hampton (2007:9) and Decock and Douven (2012:5).

Figure 5 the horizontal axis represents the distance from the prototype (measured in per cent of the distance to the Voronoi border), while the vertical axis shows how much of the influence has been lost in the point in question (measured in per cent). It is, of course, not difficult to construct other, alternative functions that capture similar intuitions. However, for the purpose of the present paper it suffices to note that one can easily capture the intuition that the distance between a non-prototypical case and the nearby prototypes affect how much influence each prototype should be allowed to have on the all-things-considered verdict. We can leave it open whether it is the function depicted in Figure 5 or some alternative function that describes this relationship.

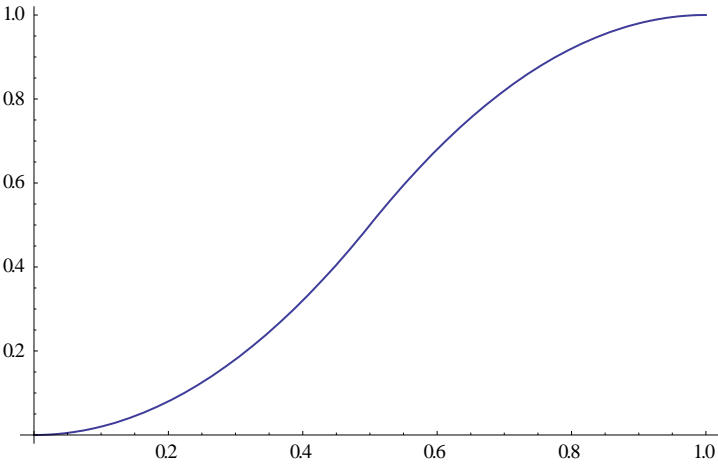


Figure 5. Hampton's S-shaped function

The further apart the prototypes for a source of normativity are, the larger will the normative grey area be, everything else being equal. Figure 6 illustrates an example in which each source of normativity is defined by an area of prototypes consisting of the infinite number of points within the square demarcated by the black dots. Only the corners of each square are visualized in the figure.

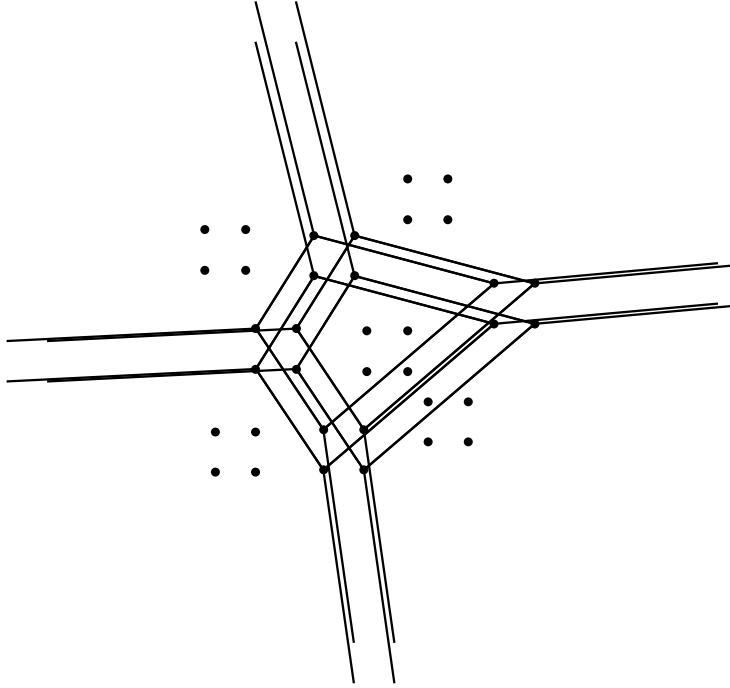


Figure 6. Large normative grey areas.

7 Four Types of Cases

The theory of radical evaluative ignorance proposed here can accommodate four types of cases. First, some cases are *prototypes* for some source of normativity. We may call these Type I cases. In Type I cases the conclusion of the normatively conscientious agent is entirely determined by the source of normativity that defines the prototype in question.

Type II cases are cases that are not prototypes for any source of normativity, but which are similar to some prototype without also being similar to a prototype for some other source of normativity. Put in a geometric vocabulary, a Type II case lies closest to only one seed point, for one region in the Voronoi diagram, no matter what seed points for other sources we compare with. In Type II cases the conclusion of the normatively conscientious agent is, just as in Type I cases, entirely determined by the source defined by the nearest prototype, without any contribution from any other source.

Furthermore, as explained in Section 3, there are also cases located exactly on the border between two or more sources or seed points. In such Type III cases the equi-distant sources contribute equally much to the all-things-considered ought.

Finally, there are also cases in which some sources located at different distances from a non-prototypical case contribute to the all-things-considered ought. These cases are ones in which some source has more than one prototype, as explained and discussed in the previous section. I will refer to these cases as Type IV cases. The relative strength of each prototype can be described by the function in Figure 5.

The all-things-considered ought in Type I and Type II cases is binary. The agent either ought to ϕ all things considered, or ought to not- ϕ all things considered. This is because in Type I and Type II cases the all-things-considered ought is entirely fixed by a sole decisive source of normativity. There is no clash between different sources. One source is always closer to the case faced by the agent no matter what other sources she compares with. But how should we analyze Type III and Type IV cases? A possible answer is that Type III and Type IV should be analyzed in non-binary terms, meaning that the all-things-considered ought is *gradual* rather than an all-or-nothing affair.

Instead of claiming that one either ought, all things considered, to ϕ or to not- ϕ , advocates of the a non-binary theory of normativity could claim that in Type III and Type IV cases, one ought to *some degree*, all things considered, to ϕ . At the same time one also ought to some (other) degree, all things considered, to not- ϕ . This claim about non-binary all-things-considered oughts is less exotic than one might think. Moral dilemmas are often thought to include cases in which moral oughts clash. You ought morally to ϕ at the same time as you ought morally to not- ϕ . What is being proposed here is that something similar sometimes hold true for the all-things-considered oughts reached by the normatively conscientious agent facing radical evaluative ignorance.

Needless to say, not all philosophers believe that moral dilemmas exist. So what argument, if any, could one give for the non-binary analysis of Type III and Type IV cases? It seems that the best argument is that all nearby prototypes carry the same normative weight as prototype cases, meaning that we should give each of the applicable prototypes their due when determining the all-things-considered ought. Consider, for instance, a Type IV case, i.e. a case located in the “overlapping” normative grey area in Figures 3 and 4. To give all prototypes their due means that all applicable prototypes matter but the relative weight of each prototype depends on the distance between the non-prototypical case under consideration and the prototype in question. That is, all applicable prototypes carry the same weight qua prototypes, meaning that they all should be given the same due in prototypical cases. However, in non-prototypical cases of Type IV, some

applicable prototypes are located further away than others. To give all prototypes their due when one faces a case located in the area between the dashed and the non-dashed line in Figure 3 therefore means that one ought to comply with more than one source of normativity, simply because there is no unique closest (most similar) source of normativity.

The reason for concluding that the all-things-considered ought comes in degrees in some cases is that this view reflects the complexity of the normative landscape better than any binary view. If there is no unique most similar prototypical case, it would arguably be too heavy handed to conclude that one source nevertheless trumps the others.

It is an open question what decision rule agents should apply if the all-things-considered ought comes in degrees. A possible rule, which I will not try to give any argument for here, is to randomize. That is, if the normatively conscientious conclusion is that all things considered you ought to some degree to do something, at the same time as you ought to some degree not do it, it seems reasonable to claim that the behavior that fits best with this complex normative landscape is to let a (possibly weighted) coin toss guide the agent's choice. This rule is, however, not the only rule worth considering. Another possibility is to always perform the act that is right to the highest degree. For extensive discussions on these rules, see Peterson (2013) and (2015).

8 Conclusion

I have demonstrated that a normatively conscientious agent can deal with radical evaluative ignorance by comparing how similar her situation is to prototypical cases for different sources of normativity. In principle, the most conscientious choice is to apply the prototypical source that is most similar to the case she is facing. However, because some sources may have more than one prototype normative gray areas may arise, in which the all-things-considered ought is, or should at least be treated as, a non-binary entity.²⁵

REFERENCES

²⁵ I would like to thank Lieven DeCock and Rik Peels for very helpful comments on previous drafts.

- Churchland, Paul M. (2007). *Neurophilosophy at Work*. Cambridge University Press.
- Copp, D. (1997). “The Ring of Gyges: Overridingness and The Unity of Reason.” *Social Philosophy and Policy* 14: 86–106.
- Decock, L. and I. Douven (2014). “What Is Graded Membership?” *Noûs* 48 (4):653-682.
- Hare, R. M. (1952). *The Language of Morals*. Oxford Clarendon Press.
- Gärdenfors, P. (2004). *Conceptual Spaces: The Geometry of Thought*. MIT Press.
- Gärdenfors, P. (2014). *The Geometry of Meaning*. MIT Press.
- Gustafsson, J. E., & Torpman, O. (2014). In Defence of My Favourite Theory. *Pacific Philosophical Quarterly*, 95(2), 159-174.
- Hampton, J.A. (2007) “Typicality, Graded Membership, and Vagueness.” *Cognitive Science* 31:355–384.
- Korsgaard, C. M. (1996). *The Sources of Normativity*. Cambridge University Press.
- Lockhart, T. (2000). *Moral Uncertainty*. Oxford University Press.
- Nietzsche, F. (2004). *Twilight of the Idols*. Dover Publications.
- Parfit, D. (2011). *On What Matters*. Oxford University Press.
- Peterson (2013). *The Dimensions of Consequentialism*. Cambridge University Press.
- Peterson (2015). “The Dimensions of Consequentialism: Reply to Schmidt, Brown, Howard-Snyder, Crisp, Tanyi and Andric, and Gertken”, *Ethical Theory and Moral Practice*, in press.

Rosch, E. (1975). "Cognitive reference points." *Cognitive Psychology*, 7(4), 532-547.

Rosch, E. H. (1973). "Natural categories." *Cognitive Psychology*, 4(3), 328-350.

Sepielli, A. (2013). "Moral Uncertainty and the Principle of Equity among Moral Theories". *Philosophy and Phenomenological Research*, 86(3), 580-589.

Sidgwick, H. (1874). *The Methods of Ethics*. Thoemmes Press.

Singer, P. (1995). *How Are We to Live?: Ethics in an Age of Self-Interest*. Oxford University Press.

Tännsjö, T. (2010). *From Reasons to Norms: On the Basic Question in Ethics*. Springer.

Tversky, A. and Gati, I. (1978). "Studies of similarity". *Cognition and categorization*, 1:79-98.

Unger, P., 1996. *Living High and Letting Die*, New York: Oxford University Press.