Reasoning with comparative moral judgements – an argument for Moral Bayesianism

Abstract

The paper discusses the notion of reasoning with comparative moral judgements (i.e. judgements of the form "act a is morally superior to act b") from the point of view of several meta-ethical positions. Using a simple formal result, it is argued that only a version of moral cognitivism that is committed to the claim that moral beliefs come in degrees can give a normatively plausible account of such reasoning. Some implications of accepting such a version of moral cognitivism are discussed.[1]

Introduction

Moral judgements[2] are an object of reasoning. People consider them, try to justify them, change them in light of new arguments and so on. This is particularly true regarding comparative moral judgements (CMJs), i.e. judgements of the form "a is morally superior to b", when "a" and "b" are either acts or states of affairs.

Comparative moral judgements, many (but not all) philosophers believe, ought to be consistent in the decision theoretic sense. Minimally, they ought to be transitive. In this paper, I argue that accepting the latter claim while also accepting the claim that agents can reason with comparative moral judgements commits one to accepting a meta-ethical position that I call moral Bayesianism (MB). Moral Bayesianism is a form of moral cognitivism, the thesis that moral judgements are beliefs. It is the version of moral cognitivism according to which moral beliefs come (just like non-moral beliefs) in degrees of credence.

The structure of the argument is as follows. I will first present a characterization of reasoning with CMJs and will argue that all those who deny MB must accept this characterization (those

---

[2]The term "moral judgements" is used in our everyday discourse, in philosophical discourse and in discourses in other disciplines (such as psychology) in two different ways. Sometimes it is used in order to refer to an act (typically a verbal act) of "judging". E.g. when I say "it is wrong to lie" I am performing a verbal act and this act is a moral judgement. Other times, however, it is used in order to refer to the mental attitude that is expressed by such an act. E.g. the reason for my saying "it is wrong to lie" is my mental judgement that it is wrong to lie. In this paper I am going to use the term in the second way.

who accept MB, however, must reject this characterization). I will then show that, under this characterization, reasoning with CMJs is impossible if we accept the transitivity of CMJs. In the last section I will show that moral Bayesians do have a good way to both characterize and conduct reasoning with CMJs.[3]

The rest of this paper will be organized in the following way. In section 1, I present MB in more detail and try to identify motivations for accepting it. The discussion in this section also serves as a good introduction to the argument that follows as it leads, in a natural way, to the problem of moral reasoning. In section 2, I present a very general formal characterization of reasoning with CMJs and argue that all non-Bayesians should accept it, while Bayesians can (and should) reject it. In section 3, within the framework of the characterization introduced in section 2, I present three axioms that whoever accepts the characterization must accept, and show that they are inconsistent with each other. In section 4, I argue that moral Bayesians have a good alternative way of characterizing reasoning with CMJs that does not lead to similar problems. This last discussion also explains what it is about non-Bayesian approaches that makes them unfit to explain reasoning with CMJs.

Moral Bayesianism

In this paper I take moral cognitivism to be the thesis that moral judgements are beliefs.[4] MB is the thesis that these beliefs come (just like non-moral beliefs) in degrees. Although the term "Moral Bayesianism" is mine, the position it denotes is not new. Many philosophers explicitly accept it and many others would be ready to accept it if it were presented to them.[5] There are, however, philosophers that explicitly deny the thesis contained in this position.

It will be useful to distinguish two types of opponents to MB. On the one hand are non-cognitivists. Non-cognitivists deny that moral judgements are beliefs and so they deny that moral judgements are quantitative beliefs.[6] On the other hand are non-Bayesian moral cognitivists. Such moral cognitivists accept that moral judgements are beliefs, but deny that they come in degrees. Moral judgements, they argue, are beliefs of a binary form: you can either hold

---

[3]The formal result presented in section 3 applies, of course, not only to CMJs, but also to comparative judgements regarding any other type of linear ordering (such as rational preferences for example). In fact a version of it can be applied to any set of propositions which is governed by a semantic that rules out as inconsistent at least one possible distribution of truth values over any three propositions in the set. However, my justification for the claim that the characterization of reasoning with comparative judgements that I present should be accepted by all non-Bayesian (with regards to a given type of comparative judgements) does not apply to reasoning with comparative judgements about all types of linear orderings. Specifically, I do not try to suggest (and I think it is not the case that) it holds with regards to a rational agent's personal preference ordering.

[4]In any case, it takes moral judgements to have truth values. For a discussion of different types of moral cognitivist positions see Van Roojen (2009). The discussion here is independent of the issues discussed there.

[5]Two examples of literatures in which at least some of the participants explicitly accept MB are the literature discussing "moral uncertainty" (for example see Lockhart [2000], Smith [2002], Jackson and Smith [2006] and Sepielli [2009]) and the literature discussing "the desire as belief thesis" (for example see Price [1989], Broome [1991], Oddie [1994], Piller [2000], and Bradley and List [2009]). I use the term "moral Bayesianism" to highlight the relevance – rarely discussed – of this thesis to moral reasoning.

[6]Van Roojen (2009) offers a good overview of different non-cognitivist views.

them or not, you can either accept them or deny them.[7] The negative part of my argument is directed against both types of non-Bayesians. Neither of these positions can account for moral reasoning with CMJs, or so I will argue.

Two clarifications are in order. First, the degrees in question here are degrees of belief, not the degrees of moral value of the objects of these beliefs or judgements. Both non-cognitivists and non-Bayesian cognitivists can (and probably do) accept that the objects of moral judgements usually do come with degrees of moral value. What they reject is that the judgements themselves come with degrees.

Moral Bayesians insist that the judgements themselves also come with degrees. On the one hand, I can be equally certain that it is wrong to lie and that it is wrong to murder, while still judging murder to be much worse than lying. On the other hand, I can believe that in case an early-stage foetus has the same moral status as an adult human being, having an abortion is as bad as murdering, but still be uncertain – to different degrees – whether an early-stage foetus has the moral status of an adult human being.

Second, up to this point I have discussed MB as a conceptual (possibly with some descriptive elements) thesis: a family of attitudes, which we call moral judgements, exists and MB is the thesis that these attitudes are beliefs. Now, however, I want to add a normative component to MB, which is that degrees of moral judgements obey the laws of probability. I am not going to argue for this claim. In the literature there are several justifications of this demand for the case of non-moral beliefs (for example see Joyce [1998]) and I see no special reason to believe that these justifications fail specifically when it comes to moral beliefs. I will not discuss this possibility here. However, most (but not all) of what I argue is independent of this assumption.

Why be a moral Bayesian? Well, the first reason is that moral judgements do intuitively come with degrees. Notice that the "feeling" of moral uncertainty usually arises in the following type of situations. When an agent discovers through reflection that she holds moral judgements that are inconsistent according to some criterion of consistency that she herself takes to be valid, she usually "feels" that she must give up on at least one of the inconsistent judgements. They are inconsistent, after all, so how can she hold all of them? However, typically, she also "feels" uncertain regarding which of the inconsistent judgements she must reject. All of them seem initially plausible (otherwise she wouldn't have held them in the first place), but they cannot all be held consistently at the same time. Which one(s), then, should she reject?

The last question "feels" like a question to which a good answer can be given. At least, it is clear that not all answers are good to the same degree. MB explains this phenomenon in a straightforward way: the "feeling" of uncertainty is just that, uncertainty. Non-Bayesian approaches would seem to face a much greater difficulty in explaining the phenomenon. Now, I agree that

---

[7]I do not know of any published work that explicitly argues against MB from a cognitivist point of view. However, in private conversations and other types of informal communication I have met the non-Bayesian cognitivist position quite often.

this argument[8] is a good yet inconclusive one, but I also think that its strength is frequently underappreciated. In sections 2 and 3 I will present a stronger argument for MB, according to which degrees of certainty are not only something we intuitively attach to moral judgements, but also an attribute of moral judgements that is necessary in order for us to account for the phenomenon of moral reasoning. Before I do so, however, I want to reflect a little more on the intuition that moral judgements come with degrees.

Consider the two versions of the trolley problem: the "side-track" version, in which it is possible to save 5 people from death by pulling a switch that moves a train to a side track on which it is going to kill only 1 person; and the "fat man" version, in which it is possible to save the 5 people by pushing a fat man from a bridge on the tracks thereby killing the fat man while stopping the train from hitting the 5 people. Now let us consider the following three claims regarding these two versions of the trolley problem.

- S – it is permissible to pull the switch in the side-track version.

- F – it is impermissible to push the man in the fat-man version.

- N – there is no morally significant difference between the two versions.

As is well-known (for example see Mikhail [2007]), about 70 percent of the population judge both S and F to be true. I know of no study that has explicitly examined people's attitudes toward N, but one of the central disagreements in the vast philosophical literature on the trolley problem concerns N. Those who argue in favour of accepting both S and F try to point to some differences between the two versions and argue that these differences are morally significant (significant enough to justify accepting both S and F). Those who argue against try to argue that these differences are not morally significant (or, if significant, not strong enough to justify accepting both S and F). In any case, it seems that everybody agrees that it is irrational to hold all three judgements at the same time.

When I come to evaluate the three claims, however, I tend to accept all three of them. Note that this is not just my intuitive reaction to the three claims being presented to me. Rather, this is a true description of my attitudes toward the three claims after deep reflection: after I read many papers on the subject, sat in many lectures on it, discussed the issue with many people whose opinions I appreciate and thought about it for many hours. At the same time, I take the three claims to be inconsistent with one another: if there is no morally relevant difference between the two versions it cannot be the case that in one of them it is permissible to save the 5 people and in the other not.

How can this be the case? Maybe, it cannot. Maybe, the problem is with my attitudes. I suspect, however, that many other people also share these judgements and, if this is so, many people would surely welcome an explanation of this pattern of judgments that does not attribute

---

[8]Introduced in a much more detailed way by Michal Smith (2002).

irrationality to all of us. MB allows us to do this. For a Bayesian, to accept that the three claims are inconsistent is to accept that the probability of the intersection of all three of them should get probability 0 in the credence function of a rational agent. In other words, for a Bayesian it follows from the inconsistency of the three claims that $c(F|SN) = 0$, where $c(.)$ is a rational agent's credence function. It does not follow from this, however, that if a rational agent believes to a high degree that S and N are true she should believe to a high degree that F is false.

There are infinitely many credence functions that respect the former condition while violating the latter. Here is one:

	Credence		1/3	1/3		1/3			
State		S	S	S	S	-S	-S	-S	-S
		F	F	-F	-F	F	F	-F	-F
		N	-N	N	-N	N	-N	N	-N

Table 1

Here, the agent believes in each one of the three propositions, S, F and N, with a relatively high probability of 2/3, while giving a probability of 0 to each state in which all three are true.

One way to view this type of situation is as a version of the famous Lottery Paradox (LP), introduced by Henry Kyborg (1961). In the LP, an agent finds herself accepting each one of the propositions included in some set (the proposition that says that ticket 1 will not win the lottery, the proposition that says that ticket 2 will not win the lottery, and so on), while rejecting their conjunction (i.e. that no ticket will win the lottery).

This is exactly what happens here. I tend to accept each one of the three claims, F, S and N, while rejecting their conjunction seems to me unavoidable.[9] Now, one of the main lessons from the LP is that there is nothing irrational in holding such degrees of belief. Given the evidence, it is not only not irrational but, arguably, the only rational thing to believe with high probability regarding each ticket that it will not win the lottery while at the same time believing with probability 1 that one of the tickets will win it. In other words, sometimes the evidence is such that it supports holding credence functions that constitute a LP. Sometimes the evidence is misleading.

Now, in the trolley problem case, as also in the case of many other moral debates,[10] most people seem to take our moral intuitions as an important type of moral evidence: having a

---

[9] An anonymous referee commented that, while in the trolley problem we can intuitively accept any conjunction of two out of the three propositions involved, the example presented in table 1 assigns to each one of these conjunctions a low probability (of 1/3). I am not sure that I share the referee's intuition (when I think about a conjunction of two out of the three propositions I become immediately aware of the fact that they entail the negation of the third), but in any case it is easy to construct an example in which each one of the three conjunctions (i.e. SF, SN and FN) gets a high probability while the conjunction of all three of them gets a low probability (though not 0 and see footnote 10). For example, this is the case when the agent assigns a credence value of 9/24 to SFN, and a credence value of 5/24 to SF-N, S-FN and –SFN.

[10] For example, consider Peter Singer's argument from "Famine, Affluence, and Morality" (1972), which has a structure almost identical to the one discussed here, as it is based on the apparent inconsistency between the following three claims: 1. It is obligatory to save a drowning child when the child is in a pond next to you, even

strong intuition that it is impermissible to push the fat man is evidence for the claim that it is impermissible to do so. In the literature one can find different explanations for why moral intuitions will tend to mislead us in many cases: they are emotionally loaded (see for example Greene [2007]), they are the product of an evolutionary process which took place in a very different environment from ours (see for example Singer [2005]), they are the product of heuristics that work well most of the time but give the wrong prescriptions in some cases (see Sunstein [2005]) and so on.

Thus, we have good reasons to suspect that our moral evidence will mislead us. So it is not surprising that we find ourselves in some cases in the kind of moral lottery paradox described above. There is nothing irrational about this. It is the rational reaction to misleading evidence.

So while non-Bayesian approaches must treat my attitudes to the two versions of the trolley problem and to other moral debates as irrational,[11] MB sees them as perfectly rational. Moreover, it seems that the explanation that MB gives to the apparent inconsistency among my attitudes succeeds in capturing exactly what is going on in my mind. This I take to be strong evidence for MB.

Notice now that while all the accounts mentioned above explain why we should expect our moral intuitions to mislead us in some cases, they do not tell us which intuition it is that is misleading us in a specific case. When we find ourselves in a "moral LP" of the sort described above we know that at least one of our intuitions misleads us, but we do not know which one it is. Neither our psychological knowledge nor our biological knowledge can tell us that. Thus, it seems perfectly reasonable to give a positive answer to the question 'do you accept this claim?' when asked regarding each one of the claims, S, F and N. At the same time, it seems reasonable to give a negative answer when asked this question regarding the conjunction of S, F and N.[12]

'How can you accept all three claims while you deny their conjunction?" the psychologist might ask his subject. The subject, then, would simply reply "you asked me for a binary judgement, but my attitudes here are not binary so I just went with the higher probability". This seems to me a reasonable reply.

Moreover, there seems nothing irrational in acting according to the same strategy (i.e. the strategy of "going with the higher probability"). I believe it is more likely than not the case that morality requires me to pull the switch in the side-track version, so I do. I believe it is more likely

---

if by doing this you will ruin your new pair of shoes. 2. It is permissible not to save dying children in far away countries, even if it is possible to do so for the cost of a new pair of shoes. 3. There are no morally significant differences between the case of the child drowning in a pond next to you and the case of dying children in far away countries.

[11]However, see Easwaran and Fitelson (2015) for an argument against the demand that a rational agent must hold a deductively consistent set of full beliefs at all times (Kyborg rejected this demand too). As noted, I do not take the discussion in this section to be a knock-down argument for MB. The aim of the discussion is merely to motivate MB. The real argument for MB is presented in the following sections. Still, it is worth mentioning that in Easwaran (forthcoming) an extension of the position presented in Eswaran and Fitelson (2015) is used in order to suggest a new subjective interpretation for probability. Thus, it seems (but I cannot say this with complete certainty) that, to the extent that the non-Bayesian can give a satisfactory account for the phenomenon described in the main text in terms of full beliefs, it is possible to view this account as a version of MB.

[12]And as mentioned in footnote 8, I am not sure about the case of the conjunction of only two out of the three propositions.

6

than not the case that morality requires me not to push the fat man in the fat-man version, so I don't. It is true that I believe with high probability that there is nothing that can justify this pattern of choices, but so what? Can you point to a principle of rationality that I have violated?

When do the problems with the "going with the higher probability" strategy start? They start when, by following it, an agent must violate a principle of consistency of choice. For example, when by following it an agent must choose intransitively. How can this happen?

The first thing to notice is that transitivity is a condition of rationality that applies both to choices and to comparative moral judgements (and other "betterness judgements"). Rationality requires us both to choose in a transitive way and to hold transitive comparative moral judgements. Although some philosophers deny one or both of these requirements, in this essay I take at least the second, i.e. transitivity of CMJs, as an assumption. By doing so I do not, however, exclude from my audience all those who deny this condition.

First, I take it that most scholars that deny that transitivity is a condition of rationality would still accept that it should be possible for agents to hold transitive CMJs. It is surely not irrational to do so, after all. Thus, meta-ethical or ethical accounts that do not allow for transitivity of CMJs must be seen as problematic in some way: they do not allow for patterns of attitudes that seem perfectly reasonable.

Second, the debate between MB and non-Bayesian approaches seems independent of the debate regarding the status of the transitivity condition. Most non-Bayesians, even ones who are willing to bite the bullet and accept that full transitivity of CMJs is impossible, will hesitate, or so I presume, to take this claim as one of the building blocks of their meta-ethical position. It should be possible to be a non-Bayesian even if you do believe in transitivity, shouldn't it?

So I am not going to justify either the transitivity of CMJs or the transitivity of choices conditions. Rather, I take them as assumptions (I only need the transitivity of choices condition for purposes of exposition, though. The argument does not depend on it. It does depend on transitivity of CMJs). But what exactly does the transitivity condition say?

Well, for a non-Bayesian the answer is obvious: it says that if one holds both the judgements that one act, A, is morally superior to another act, B, and that B is morally superior to a third act, C, then one ought to hold the judgement that A is morally superior to C.[13] But for a Bayesian the requirement says something else. It says that one ought to give credence 0 to any intransitive ranking of the three acts, A, B and C. It does not say that if one believes with high probability that A is morally superior to B and that B is morally superior to C then one ought to believe with high probability that A is morally superior to C. There are infinitely many credence functions that respect the former condition and violate the second. Here is one for example:

_____

[13]There are actually three possible readings of the requirement. Let "i>j" stand for "i is morally superior to j". Then, the three readings are the following ones. 1. If one accepts "A>B" and "B>C" then one ought to accept "A>C". 2. If one accepts "A>B" then one ought to accept "if 'B>C' then 'A>C'". 3. One ought to accept "if 'A>B' and 'B>C' then 'A>C'". I think one should accept the requirement under all three readings. In any case, nothing in my argument depends on this.

1/3	1/3	1/3
A	C	B
B	A	C
C	B	A

Table 2

Here, the agent believes with probability 2/3 in any one of the three CMJs, "A>B", "B>C, and "C>A", but still gives credence 0 to the possibility that the moral superiority relation is intransitive.

However, when one holds such degrees of beliefs then, by following the "going with the higher probability" strategy, one is led to choose in an intransitive way. This is something one wants to avoid, I assume. What should an agent that finds herself in such an epistemic condition do?

Well, there are two different questions here. The first question is how should she choose? I am not going to discuss this question here (I discuss it in Nissan-Rozen (2012)). Most of the "moral uncertainty" literature deals with exactly this question. However, there is another question that does not receive enough attention in the literature and arises in this context. This is the question, not of choice, but of reasoning: when the agent finds herself in such an epistemic condition she knows for sure that at least one of the claims in which she believes with high probability is false. Thus, if there is no need for an immediate choice, it is clear that the agent ought to try and reason herself out of the unfortunate epistemic position she has found herself in. She can, for example, try to collect more moral evidence (consult with people whose opinions she values, think of more arguments, read some books, etc.), or revaluate the evidence she already has.

Notice that the same holds also for non-Bayesians. When a non-Bayesian agent finds himself accepting inconsistent CMJs, and when there is no need for an immediate decision, he too ought to try and reason himself out of the inconsistency. In the next section I am going to step out of the Bayesian point of view, at least for the time being, and will attack the problem of reasoning with CMJs from the non-Bayesian point of view. I will return to the Bayesian point of view only after we reach a dead-end by applying non-Bayesian thinking to the problem.

Reasoning with CMJs

At the most general level, reasoning with CMJs is a process in which an agent changes some of the CMJs he holds in a non-arbitrary way. It is true, of course, that reasoning might not lead an agent to change any one of the CMJs he holds, but – if we want to discuss a non-trivial concept of reasoning – we must also allow for cases in which reasoning does lead the agent to change some of the CMJs he holds. Moreover, in the kind of cases that motivates my discussion of reasoning,

cases in which an agent realizes through reflection that he holds inconsistent CMJs (according to a criterion of consistency he accepts), the reasoning must lead the agent to change at least one of the CMJs he originally held. Otherwise, he would end the process of reasoning holding inconsistent CMJs, which we can assume cannot be the case for a rational agent.

Of course, not any process in which an agent changes some of the CMJs he holds constitutes a process of "reasoning". In order for an instance of a change in one's CMJs to constitute an instance of reasoning with CMJs the change should be due to a reason. An arbitrary change is not a matter of reasoning.[14] I will exploit this last distinction shortly in order to present a necessary condition for an instance of a change in one's CMJs to constitute an instance of reasoning with CMJs. Before I do so, however, one comment must be made.

There might be (there probably are) some non-cognitivists who will not mind accepting that reasoning with CMJs is impossible (some Humean non-cognitivists will probably even be happy to accept this claim). Such non-cognitivists take a change in one's CMJs to be a similar matter to a change in one's preferences between, say, vanilla ice cream and chocolate ice cream. My argument is not directed at such non-cognitivists. My argument takes as an assumption that reasoning with CMJs is possible and, using this assumption, argues for MB. It does not assume, however, that reasoning must be made with quantitative beliefs. In my starting point I allow for both reasoning with qualitative beliefs and with other attitudes (for discussions regarding the idea of reasoning with preferences see, for example, Broome [2006]).

Although some non-cognitivists will be willing to accept that reasoning with CMJs is impossible, most non-cognitivists, I believe, will not. Most non-cognitivists do want to accommodate within their meta-ethical framework most of the features people intuitively attribute to morality. The idea that it is possible to reason with moral judgements – in the minimal sense that I have presented – is surely one of these features.

Those who do believe that reasoning with CMJs is possible must accept the very general characterization I have presented here for such reasoning. I am now going to use this characterization in order to present a necessary condition for an instance of a change in one's CMJs to constitute an instance of reasoning with CMJs. In order to do so we have to distinguish, first, between two stages involved in a process of (at least some instances of) reasoning with CMJs.

The first stage is the stage in which the agent – after being exposed to some information he takes to be morally significant (I will call this "moral evidence") – decides to change at least one of the CMJs he holds. The second stage is the stage in which, following this change, and as a result

---

[14]I am not trying to perform here a conceptual analysis of the concept of reasoning. Maybe an arbitrary change in one's CMJs does constitute an instance of "reasoning" with CMJs. John Broome's (2006) discussion of reasoning with preferences seems to assume that this might be the case. Broome, then, limits his discussion to what he calls "correct reasoning". I do not mind, of course, accepting such terminology. The important point is that there is a sense in which we do demand from a rational moral agent, who is involved in a process in which he changes some of the CMJs he holds, to do so in a non-arbitrary way. Whether this demand is a necessary condition in order for the process to count as "reasoning" or only a necessary condition in order for the process to count as "correct reasoning" does not concern me. For convenience, then, I will assume that the former condition holds.

of it, the agent changes other CMJs he holds in order to keep the set of judgements he holds consistent.

For example, consider an agent who judges one act, A, to be morally superior to another act, B. Since he judges A to be superior to B, he also judges, we can assume, that a lottery that brings A with probability 0.9 and B with probability 0.1 is morally superior to a lottery that brings B with probability 0.9 and A with probability 0.1. The agent is then exposed to a piece of moral evidence that makes him change his judgement regarding the moral superiority relation that holds between A and B, and judges now B to be morally superior to A. Maybe, for example, he thinks of an argument he had never thought of before, or maybe he reads a book with such an argument, or maybe he does not think of any argument but simply learns that a person, whose opinion he values, judges B to be superior to A, or maybe he has a dream in which God tells him that B is superior to A. I do not assume anything, at this stage, about what can justifiably constitute moral evidence (later on I will present one weak assumption about what cannot constitute moral evidence).

In any case, the agent changes his judgement regarding the moral superiority relation between A and B, and this change brings with it other changes in the set of CMJs the agent holds. For example, since the agent now judges B to be superior to A, he should also change his judgement regarding the moral superiority relation that holds between the lottery that brings B with probability 0.9 and A with probability 0.1 and the lottery that brings A with probability 0.9 and B with probability 0.1. He should now judge the former to be superior to the latter.

The point of this example is not to argue that the agent should normatively change his judgements in such a way. I believe he should, but even if he shouldn't there must be other cases in which an initial change in one's CMJs carries with it further changes. The point is rather to make clear the distinction between the two stages involved in the process of reasoning with CMJs: the stage in which an agent – after being exposed to some moral information – performs one change in the set of CMJs he holds, and the stage in which the initial change brings some further changes.

Now, the condition I will shortly present applies only to the second stage, not to the first stage. There might be specific instances of reasoning with CMJs in which the initial change does not bring with it any further changes because it conflicts with none of the other CMJs the agent holds, but if we want to allow for non-trivial reasoning we must also allow for cases in which the initial change does bring with it some other changes.

In the same way, there might be some instances of reasoning with CMJs in which the second change I was referring to is absent because the agent's immediate response to the moral evidence he is exposed to is a change in each one of the CMJs he holds (and thus there is no need for the second stage in which the agent adjusts the set of his CMJs to the initial change). However, arguing that there are no cases of reasoning with CMJs in which the second stage plays a significant role trivializes the concept of reasoning used. Intuitively, most instances of reasoning

with CMJs are instances in which the evidence the agent is being exposed to (e.g. an intuition triggered by an "intuition pump") supports a change in a set of CMJs that is a proper subset of the set of all CMJs the agent has.

Thus, accounting for the second stage I have described is necessary for anyone who wants to allow for a non-trivial concept of reasoning with CMJs.

Here is the idea: when one reasons with one's CMJs one changes them in a way that is only sensitive to what one takes to be relevant considerations. To be clear, the demand is not that the way one changes one's CMJs will always be sensitive to what one takes to be relevant considerations. This demand is too strong. Sometimes agents judge some considerations to be relevant to their CMJs but, nevertheless, do not take them to be strong enough to trigger a change in the CMJs they accept. The demand I have in mind is a much weaker one. It is the demand that the way one changes one's CMJs is not sensitive to what one takes to be irrelevant considerations.

Consider an agent who changes her CMJs and is then asked to supply a reason for the change. The agent can either deny that there was any reason for the change or supply a reason. In the first case, the change is surely not an instance of reasoning with CMJs as the agent herself claims there was no reason for the change. In the second case, when the agent does supply a reason for the change, if she does so in a way that violates the condition then the reason she supplies must be a reason that the agent herself takes to be irrelevant to the change she has made. This just means that it is not a genuine reason for the change.[15]

We are thus looking on a process in which, following an initial change in one's CMJs, one changes other CMJs one previously held and, while doing so, one takes the initial change to be relevant to the other changes one makes. If you are a non-Bayesian, you can think of such a process in the following way. There is a function, f(.), which takes as input ordered pairs, in which the first element is a ranking of alternatives that belong to some set of alternatives, A, and the second element is an order of all the alternatives in some subset of A, and produces as output an order of all the alternatives in A.

The interpretation is quite simple. An agent starts with some set of CMJs regarding some of the alternatives available to her. These need not be complete or transitive, but we can assume that they are as this does not matter for the discussion that follows and it will make the presentation simpler. Then, she receives some input. This can be the result of either the agent thinking about the matter, or discussing the issue with other people, or reading a book, or receiving a message from God, or what have you. The input, however, must be in the form of an order (i.e. a complete and transitive relation) over some of the alternatives available to the agent. In the

---

[15]You can read the last two paragraphs while interpreting the phrases "takes to be a reason" and "takes to be relevant" anyway you like. I think of them as referring to beliefs: "takes X to be a reason" means believe that X is a reason; but if you prefer to understand the agent's attitudes toward the status of the information he gets differently this is okay too. In any case, if you do want to talk about reasoning in the way I have characterized here, you must make a distinction between information that the agent takes to be relevant and information that he does not.

simplest case – which is sufficient for us to consider (since problems begin even with it) – the input is just a moral superiority relation between two of the alternatives.

Now, the agent must decide how to incorporate the input she has received into her set of CMJs. I assume that the agent will do so in such a way that the result will be an ordering of the alternatives (this represents the agent's commitment to rationality). For example, consider an agent that holds the following three judgements: act a is morally superior to act b, act b is morally superior to act c, act a is morally superior to act c. Suppose the agent – following a discussion with her philosophy professor – changes her judgement regarding the moral superiority relation that holds between a and b, and judges now b to be morally superior to a. The function that we are considering tells the agent how this should affect her other judgements.

This is a very general characterization of the process of reasoning with CMJs.[16] It assumes nothing about what triggered the change; it allows the agent to start with any kind of CMJs and is very pluralistic regarding the richness of the input. What is not allowed by this characterization is that either the input or the output will be in the form of a probability distribution over the set of all possible orders of the alternatives, or of a new probability value for a specific moral superiority relation between two alternatives (as by allowing that, one becomes a Bayesian).

Maybe it will be useful to make explicit a distinction that I have implicitly used in this characterization. This is the distinction between the information that the agent has received and the effect of this information on the agent's mental states. I assume nothing about the information itself. Different people can take different things as constituting for them moral information. Assuming that an agent did receive information that she takes to be morally relevant, and assuming that this information was strong enough to change at least one of the agent's judgements, the question of how this change should effect the agent's other judgements arises.

As explained, the characterization presented here tries to capture only this last stage – the stage in which one change in an agent's judgements brings other changes. Now, all non-Bayesians must accept that the change must take a binary form, that is, the agent must incorporate a new judgement into her set of judgements. It is true that both non-Bayesian cognitivists and non-cognitivists can argue that the information itself might support a specific change in the agent's judgements to different degrees. However, once a piece of information is powerful enough to trigger a change, the change is binary; it does not come in degrees.

Only the Bayesian can reject that. For the Bayesian, the input itself (not only the strength of the information) can come in degrees. A piece of morally relevant information can, for the Bayesian, make an agent raise the probability she attaches to a specific judgement, without accepting it. More importantly, the judgements themselves – the outputs of the function – are not binary; they always come with a probability value attached to them. This is exactly the difference between Bayesians and non-Bayesian: Bayesians believes that beliefs – not only the evidential support

---

[16]See Hansson 1995 and 2001 (for example) for a similar characterization of reasoning with preferences and for preference change generally.

for a given belief, but the belief itself – come in degrees; non-Bayesians deny that. Thus, while Bayesians can (and should) reject the characterization presented here for reasoning (or merely revising) CMJs, non-Bayesians have no reason to do so.

In the next section I will introduce three axioms on f(.) and will argue that all non-Bayesians must accept them. I will then show that these axioms are inconsistent with each other.

Non-Bayesian reasoning with CMJs is impossible

In order to get our result it is enough for us to consider only the simplest case in which there are only three alternatives and the agent starts with an order over them and gets an input that is a relation over two of the three. Even in this simplest case, we get an impossibility result.

Let f(.) be our function, let A = {a, b ,c} be a set of alternatives, let > be a complete and transitive relation over A. We will use the notation "i>j>k" to refer to the set of CMJs that ranks i over j, j over k and i over k, and in the same way, the notation "i>j" will denote the CMJ that ranks i over j.[17] Consider now the following three restrictions on f(.):

1. Stability: if the input "l>m" is consistent with "i>j>k", then f("i>j>k", "l>m") = "i>j>k".

2. Respectfulness: f("i>j>k", "l>m") is consistent with "l>m".

3. Commutativity: for any two consistent inputs, "l>m" and "z>x", f(f("i>j>k", "l>m"), "z>x") = f(f("i>j>k", "z>x"), "l>m").

The first two axioms are necessary in order for us to treat the function as a function that describes a CMJs change (not specifically reasoning with CMJs). They are, in a sense, pre-conditions for any discussion regarding CMJs change.[18] "Stability" requires that if the agent gets no new information, i.e. if he gets an input that is consistent with his CMJs, he does not change his CMJs. "Respectfulness" requires that if the agent learns something, i.e. if he receives an input that is inconsistent with his CMJs, he incorporates it into his new set of CMJs.

The commutativity axiom is the one that characterizes the function as a function that describes reasoning with CMJs and not just any kind of CMJs change. I discuss the axiom at length in section 4. For now, I only want to point to the general motivation for accepting it. When one reasons with one's CMJs one changes one's CMJs in a way that is only sensitive to what one judges to be relevant to the question at hand. One thing that most people in most cases take to be irrelevant to the question of which CMJs they ought to have is the order in which they receive different consistent pieces of relevant information.

It does not seem to matter whether I first thought of one argument, A, that made me change one of my judgements, and then of another argument, B, which made me change another judgement

---

[17]There is an implicit semantic in the background that – for the sake of simplicity – I choose not to make explicit. It is easy to see what it should look like, though, in our simple case. Hansson 1995 and 2001 provides a much more general framework.

[18]"Stability" is a special case of Hansson's "vacuity" and "respectfulness" is equivalent to Hansson's "success".

(consistent with the judgement produced by A), or first thought of B and then of A. Of course, if the judgement produced by A is inconsistent with the judgement produced by B, the order does matter (for example, if I first learn A and then "not A" I should believe "not A, but if the order is reversed, I should believe A), but "commutativity" refers only to consistent inputs. When it comes to consistent inputs, the order does not matter, unless one holds very peculiar meta-ethical or epistemological convictions.

For example, one might insist that the order in which an agent gets different pieces of information has a constitutive role for some moral facts, or one might insist that this order – in itself – reveals to the agent something about morality (maybe God communicates with us through the order in which She sends us different pieces of information, for instance). It is hard to think of any example in which the order does intuitively seem to matter. However, even if there are some contexts like that, surely in most cases the order does not matter; yet the result that will be presented now shows that it should always matter!

Theorem: there is no function that respects "stability", "respectfulness" and "commutativity".[19]

Proof: consider WLOG, an agent that starts with a set of CMJs according to which "a>b>c". Consider first the two consistent inputs, "a>b" and "c>a". From "respectfulness" and "stability" we immediately get that f("a>b>c", "c>a") = f(f("a>b>c", "a>b"), "c>a") is equal to either "c>a>b", or to "c>b>a" or to "b>c>a". From applying "respectfulness" and "commutativity" again we get f(f("a>b>c", "c>a"), "a>b") = "c>a>b" and from "commutativity" we get f(f("a>b>c", "a>b"), "c>a") = "c>a>b". Thus, f("a>b>c", "c>a") = "c>a>b".

However, consider now the two consistent inputs, "b>c" and "c>a". Similar considerations to the ones we used in the previous paragraph show that f("a>b>c", "c>a") = f(f("a>b>c", "b>c"), "c>a") = f(f("a>b>c", "c>a"), "b>c") = "b>c>a", contrary to our conclusion that , f("a>b>c", "c>a") = "c>a>b". Q.E.D.[20]

The following table demonstrates more clearly what is going on:

[19]I was not able to locate an earlier proof of this (very simple) result. However, there are obvious formal connections between this result and the literature dealing with judgement aggregation on the one hand (for a good introduction see List 2012) and the literature dealing with preference change on the other hand (such as Hansson 1995 and 2001). The literature that explores the relation between Bayesian updating and the AGM approach to beliefs revisions is also of obvious relevancy here. See for example Lepage and Morgan (2011).

[20]It is easy to see that the result can be generalized to any system of propositions that includes at least three inconsistent propositions. The inconsistency need not be always in the form of intransitivity. I thank Christian List for pointing this out to me. However, as indicated in footnote 2, I do not try to suggest that the result is of any philosophical interest when considered under different interpretations. Maybe it is under some interpretations, but I could not think of any other than the CMJs interpretation presented here.

		A	B	C	D	E	F
		a>b	b>a	a>c	c>a	b>c	c>b
1	a>b>c				c>a>b (b>c>a)		
2	a>c>b						
3	b>a>c						
4	b>c>a						
5	c>a>b						
6	c>b>a						

Table 3

In this table columns A to F represent all the possible inputs the agent in our simple case can get. Rows 1 to 6 represent all the possible orders the agent can start from. The function we are looking for assigns to every cell in the table one of the possible orders.

We start from row 1 and consider first the (consistent) inputs "a>b" and "c>a". If the agent get the input "a>b" first, he stays in row 1 (from "stability"), and then when he gets the input "c>a" he moves to one of the rows in which c is preferred to a, i.e. to either 6 or 5 or 4. If the agent gets the input "c>a" first, he moves to either row 6 or 5 or 4, and then, when he gets the input "a>b", he must move to a row in which a is ranked above b. However, from "commutativity" we know that he must end up in either row 6 or 5 or 4. The only row out of these three in which a is preferred to b is 5. This is the reason why I wrote in cell D1 "c>a>b". However, following exactly the same line of reasoning regarding the inputs "b>c" and "c>a" leads to the conclusion that the value of cell D1 should be "b>c>a". However, the function cannot give two different values to the same argument.

The result shows that, after she receives some information, a non-Bayesian (cognitivist or not) cannot change her CMJs in a commutative way. The non-Bayesian has now two options. She can either argue that the order in which an agent receives the information always matters (either epistemically or metaphysically) or accept that indeed reasoning with CMJs is impossible. There is no other option. If she refuses to accept that the order in which different pieces of information are revealed to an agent matters, she must admit that the agent has no way to change her CMJs in a way that is not sensitive to irrelevant features of reality. If she insists that there must be such a way, she must accept that the order matters. But, of course, the order does not matter. Thus, a non-Bayesian must admit that reasoning with CMJs is impossible.

Bayesians, however, reject the claim – implicit in the formalization – that one's attitudes regarding comparative moral claims can always be represented in the form of a relation between some alternatives. For the Bayesian, the information can lead the agent to change his degrees of belief in some of his judgements, without making him change the judgements he accepts (or "believes", as the word is used in a qualitative way).

It would be misleading to argue that the Bayesian is free to reject both "respectfulness" and "stability", as what the Bayesian really rejects is the mere characterization of reasoning presented. For a Bayesian, "respectfulness" and "stability" do not make any sense: they assume that one's attitudes regarding comparative moral claims are binary, when these attitudes are not. For a Bayesian, an argument that supports a judgement an agent already holds (however you choose to cash out the term "hold" – maybe in terms of high degree of belief, for example) can make the agent raise his degree of belief in this judgement, and an argument that supports a judgement an agent rejects can make him lower his degree of belief in the judgement. This does not mean, however, that the agent changes the judgements he accepts. When it comes to reasoning, "acceptance" has no normative role for the Bayesian.

Again, this move is unavailable both to the non-cognitivist and to the non-Bayesian cognitivist. Non-Bayesians can argue that a piece of information is not strong enough to change the judgements an agent accepts, but once a piece of information is strong enough to do so, the change is binary.

Some non-cognitivists might consider adding a measure of confidence to the judgements one accepts, but still argue that this measure is not a measure of strength of beliefs. If this measure is probabilistic, then the difference between such non-cognitivists and Bayesians becomes merely terminological: comparative moral judgements behave, according to both accounts, in the same way; they just get different tags attached to them. If this measure is not probabilistic, then I am tempted to look at this move as an unorthodox version of Bayesianism, which accepts that beliefs come in degrees but rejects that these degrees should (normatively) obey the probability axioms. One can find in the literature some arguments against such a move but also some arguments in favour of it (see Joyce [1998] for a discussion).

The lesson, therefore, seems to be the following. Moral reasoning requires degrees of confidence in one's comparative moral judgements. Bayesianism allows for this. It takes degrees of confidence in comparative moral judgements to be degrees of belief in the propositions that are the objects of the judgements. All versions of non-Bayesianism do not allow for this and thus they cannot give a normatively plausible account of reasoning with comparative moral judgements.

The fact that Bayesianism is not vulnerable to the result presented in this section does not mean, however, that it is on safe ground. A plausible account of Bayesian reasoning with comparative moral judgements should still be presented. In the next section I will argue that such an account is, in fact, available. The discussion will also make clear what it is about non-Bayesianism that makes it unfit to account for moral reasoning.

What can the Bayesian do?

The first thing to notice is that, for the case in which the input itself is learned with certainty (although the agent can still be uncertain regarding other judgements), classical Bayesian updating is a commutative method of reasoning that has many other virtues that are well discussed in the literature.

However, especially when it comes to moral judgements, it seems that we must allow for probabilistic inputs: in most cases the morally relevant information we get does not make us absolutely certain in a specific moral judgement; rather it makes us raise the probability we attach to it. In such cases, the standard Bayesian move is to use Jeffrey's conditionalization, which allows for belief updating when nothing is learnt with certainty (see Jeffrey (1992)).

Bayesian updating can take place only if the agent has become completely certain that some proposition is true; that is; if he has raised the probability he assigns to some proposition to 1. Jeffrey (1992) argued that not all cases of learning can be characterized in such a way.

In cases where this happens, however, we can still demand that the way the agent updates his beliefs will obey some consistency conditions. Jeffrey claimed that the main consistency condition the agent should obey is that, after raising the probability of some proposition, A, the conditional probability of all other propositions given A should stay the same. He showed that this condition is satisfied iff the following condition holds for every proposition, B$\neq$A:[21]

Jeffrey's Conditionalization (for the simple case): if p(A) $\neq$ O, 1,

q(B) = q(A)p(B|A) + q(¬A)p(B|¬A)

Here, p(.) is the agent's initial probability distribution and q(.) is his new probability distribution after raising the probability of A. It is easy to see that classical Bayesian updating is just a special case of JC.

Jeffrey's conditionalization is commutative, however, only in special cases. It is commutative for two inputs when the two inputs are probabilistically independent. When two inputs are probabilistically dependent, the order in which an agent considers the two inputs does matter for the final probability distribution the agent holds.

Does this feature of JC constitute a problem for MB? The answer is, at least on the face of it, yes. To see this, consider the case in which there are 6 states and each state is one of the possible orders over three alternatives, A, B and C.

Question 1: is there a probability distribution over the 6 states such that each state gets a positive probability and all the events of the form ("X>Y") are probabilistically independent of each other?

---

[21] For simplicity I will deal here only with the simple case in which the agent has only raised the probability he assigns to one proposition. Nothing in the discussion depends on this.

Answer: no.

Question 2: is there a probability distribution over the 6 states such that every state gets a positive probability and two pairs of such events (e.g. "A>B" and "A>C" and "B>C" and "A>C") are probabilistically independent?

Answer: no.[22]

There is, however, a probability distribution such that only 4 states get positive probability and there are two pairs of such events which are probabilistically independent. Here is one:

A	A	B	B	C	C
B	C	A	C	A	B
C	B	C	A	B	A
	1/16	3/16	9/16	3/16	

Table 4

Here both ("B>C" and "C>A") and ("A>B" and "C>A") are probabilistically independent:

p(C>A) = 12/16

p(A>B) = 4/16 = 1/4

p(A>B|C>A) = p(C>A>B)/ p(C>A) = 3/12 = $\frac{1}{4}$

and

p(C>A) = 12/16

p(B>C) = 12/16 = 3/4

p(B>C|C>A) = p(B>C>A)/ p(C>A) = 9/12 = $\frac{3}{4}$

Now, since both "B>C" and "A>B" are probabilistically independent of "A>C", Jeffrey's conditioning on both pairs ("B>C" and "C>A") and ("A>B" and "C>A") is commutative.

This happens, however, only when the order "A>B>C" gets probability 0. If "A>B>C" gets a positive probability, then Jeffrey's conditioning is not commutative over at least one of these pairs.

Notice that this phenomenon is analogous to the one that caused the problems in the non-Bayesian framework: when one starts from A>B>C, f(.) cannot be commutative (if the two other axioms are respected) over both ("A>B" and "C>A") and ("B>C" and "C>A"). In the Bayesian case our conclusion is that Jeffrey's conditioning is commutative with regards to both ("A>B" and "C>A") and ("B>C" and "C>A") only when p(A>B>C) = 0.

---

[22]The proofs are straightforward but still they require some space. As the questioner's purpose is only to highlight how severe is the threat that JC's incommutativity poses to MB (which is the position I argue for), my argument in this paper does not depend on the truth of the answers I gave in the main text and so, for the sake of simplicity, I chose to omit the proofs. The answers are correct, though.

So, on the face of it, it seems that Bayesians face an analogous problem to the one non-Bayesians face in accounting for moral reasoning. I will argue shortly that this is not the case, but before I do so that it might be useful to discuss why it is that JC is not commutative (in those cases that it is not).

One way to understand the incommutativity of JC is the following.[23] When two inputs are probabilistically dependent then, when an agent learns one of them, she does not only learn something about the proposition that is the object of the judgement in question, but also about how she ought to evaluate the other input in case she receives it.

For example, consider an agent that starts with a probability distribution according to which the probability of a>b depends on the probability of c>a, so p(a>b|c>a) ≠ p(a>b). Now the agent gets some information that makes her raise p(c>a), but still keep it below 1. If the agent uses Jeffrey's conditionalization, one effect of receiving this input is that the agent will change p(a>b), but this is not the only effect. The agent must also change now p(c>a|a>b). In other words, the agent must change the way she evaluates c>a under the supposition that a>b is true. Thus, she learns something not only about the moral superiority relation that holds between a and c, and not only about the moral superiority relation that holds between a and b, but also about what is the right way to evaluate inputs she might receive in the future.[24]

In such a case, it can be argued, commutativity ought to be violated because the method of reasoning has changed: after the agent learned the new input she evaluates moral information in a different way and thus different orders should lead to different outcomes. But if this is so, then surely there is a sense in which the information that the agent gets when she learns, first the input that makes her change the probability she attaches to a>b, and then the input that makes her change the probability she attaches to c>a, is different from the information the agent gets when she gets the inputs in the opposite order.

This, then, should lead us to question the justifiability of the commutativity axiom: what we just saw is that there is an intuitive sense in which the order in which an agent gets two consistent pieces of information can matter epistemically. It matters epistemically when it teaches the agent something about the way she would update her beliefs in a possible future. We also saw that in the kind of cases we are interested in, this is always the case!

Can this be true? Can it be true that the order in which an agent gets any two possible consistent pieces of moral evidence always teaches the agent something about the way she would update her beliefs in the future? Of course not. For example, if the agent gets two inputs that make her raise the probability she attaches to both a>b and b>c to 1, the order does not teach her anything about possible future belief updating: no matter which of the two inputs the agent gets first, she would end up believing with probability 1 in a>b>c, and this belief will not change

---

[23]I thank Richard Bardley for explaining this point to me.

[24]Now we can see that the case of two inconsistent inputs is just a limiting case of two probabilistically dependent inputs.

after gaining any new piece of evidence.

So what is going on here? Why is it the case that when dealing with learning certainties Bayesian updating (which is a special case of JC) is commutative while when dealing with learning without certainty JC is not commutative? The answer is simple and is not new to Bayesians. It will, however, be helpful to consider it here in the context of our inquiry.

When it comes to learning certainties, the evidence the agent learns can be identified with one of the propositions in his algebra. To say that an agent raised the probability he attaches to a>b to 1 is to say that the agent has learnt a>b. However, when it comes to learning without certainty, this is not the case. The agent was exposed to a new piece of evidence that cannot be captured neatly by any one of the propositions in his algebra. This is exactly why he does not raise the probability of any proposition in his algebra to 1. So we must make a distinction between the evidence to which the agent was exposed and the effect of this evidence on the agent's degrees of belief in the propositions in his algebra (in our case, the propositions, a>b, b>c, a>c, their conjunctions, disjunctions, and negations). It will be convenient to use the term "input" (which I have already used in this way) in order to refer to this effect.

Using the above distinction we can make another distinction, a distinction between two conditions of commutativity. The first condition is commutativity with respect to inputs (i.e. the effect of the evidence on the agent's degrees of belief). The second condition is commutativity with respect to experiences (i.e. the different pieces of information to which the agent was exposed).

Now, it is clear that the incommutatitivty of JC we have discussed is incommutativity of the first type, not the second. But of course, this second kind of incommutativity should not be ruled out: raising the probability of a proposition from 0.2 to 0.4, and then from 0.4 to 0.6, must, of course, lead to a different probability distribution than raising the probability of the same proposition from 0.2 to 0.6 and then lowering it from 0.6 to 0.4.

We want the way in which we update our beliefs to be commutative with respect to experiences, not with respect to their effects. The rationale for commutativity is the following. Holding your initial belief state constant, If you are exposed to the same pieces of evidence you should end up in the same belief state and the order in which you are exposed to different pieces of evidence is not a piece of evidence in itself. This does not mean, however, that the way we update our beliefs should be commutative with respect to the effect of the evidence.

To make things clearer consider again the example above while keeping in mind that there is an underlying stream of evidence that is responsible for the changes in the agent's degrees of belief. In the example an agent raises the probability of a proposition from 0.2 to 0.4 after being exposed to an experience, x, and then raises it from 0.4 to 0.6 after being exposed to another experience, y. The same agent would raise the probability of the same proposition from 0.2 to 0.6 after being exposed to an experience, y', and would lower the probability of this proposition from 0.6 to 0.4, after being exposed to an experience, x'.

In order for the incommutativity with respect to the inputs to be also incommutativity with respect to the experiences it must be the case that x=x' and y=y', but this is surely not the case. While x made the agent raise the probability of the proposition, x' made the agent lower this probability. While y made the agent raise the probability of the proposition from 0.4 to 0.6, y' made him raise this probability much more, from 0.2 to 0.6.

So, generally speaking, incommutativity with respect to inputs does not automatically translate into incommutativity with respect to experiences. Is JC commutative with respect to experiences? Well, the answer depends on the type of experiences the agent is being exposed to. Hartry Flied (1978) showed that there is a large family of experiences regarding which JC is commutative. This family is the family of all experiences that we can identify using Bayes factors, i.e. using the expression q(A)p(-A)/q(-A)p(A).

For example, an agent who starts with degrees of beliefs that can be represented by a probability distribution p(.) is exposed to a piece of evidence, X, that makes him change his degree of belief in a proposition, A, in such a way that his new degrees of beliefs can be represented by the probability distribution q(.). We can identify X using the Bayes factor we calculate using the change in the agent's degrees of belief in A. in the second stage, we can deduce what will be the effect of learning X on the agent's degrees of belief if he had started from a different degree of belief in A.

Here is a short demonstration that might help to show things more clearly. Assume the agent starts from a probability distribution p'(.) and is now exposed to X, what should be his new degree of belief in A (in his new probability distribution, q'(.))?

The answer is simple. By the assumption that we can identify X with the Bayes factor, we know that:

X = q'(A)p'(-A)/q'(-A)p'(A)

Since we know the value of X (calculated using p(A) and q(A)) and the value of p'(A), we can find out the value of q'(A).

So we can point to a well-defined family of experiences regarding which JC is commutative. It turns out (it is easy to see this by looking at q(A) as a function of p(A) while keeping the Bayes factor constant) that these experiences are experiences that have a decreasing marginal contribution to the agent's degree of belief, i.e. the same experience makes the agent raise the probability of a proposition less the higher is his initial degree of belief in the proposition (this feature does not fully characterize the family but it is a feature of all its members). Since our intuitions regarding these matters are not very vivid, I leave it to readers to think of what they take to be paradigmatic cases of moral evidence and see if these cases have this property. I believe they do.

There is one last worry that should be dealt with. I have just argued that commutativity with

respect to the effect of experiences is not the kind of commutativity we should demand from an updating method. Instead we should demand commutativity with respect to the experiences themselves. While JC is commutative (for a large family of experiences) in the latter sense, it is not commutative in the former sense. Can the same type of argument work for the non-Bayesian?

The answer is no, and it will be revealing to show why this is the case. Notice that by replacing my original commutativity condition for the non-Bayesian characterization of reasoning with a commutativity condition with respect to experiences we can only exchange the impossibility result for a triviality result.

To see this recall that the result discussed in section 2 shows that f(.) cannot be commutative at the same time regarding both (C>A and A>B) and (C>A and B>C). Let us assume it is not commutative regarding (C>A and A>B). Now suppose the agent's reasoning is commutative with respect to experiences. Thus, there can be no two experiences, x, y, such that x makes the agent change B>A to A>B and stay on A>B if the agent is there and y makes him change A>C to C>A and stay on C>A if the agent is there. If there were two such experiences, f(.) would be commutative with regards to (C>A and A>B), which we assumed is not the case. Thus, it is either the case that:

1. No experience is powerful enough to change one of these judgements.

Or

2. The same experience can change B>A to A>B and A>B to B>A.

Option 1 trivializes reasoning. Accepting "1" means accepting that out of every possible three alternatives, there is at least one pair of alternatives such that the agent's judgement regarding the moral superiority relation that holds between them is such that nothing can make the agent change it.

Whether option 2 also trivializes reasoning is a matter of interpretation, but it surely significantly limits the scope of instances of reasoning that remain open for the non-Bayesian. Accepting "2" means that out of every possible three alternatives there is at least one pair of alternatives such that the agent's judgement regarding the moral superiority relation that holds between them is such that only experiences that behave in a very special way can make the agent change them: experiences that make the agent reject the judgement that he holds (regarding the superiority relation between these two alternatives), no matter what it is.

Arguably some types of high-order evidence, such as pieces of evidence that undermine the evidential relevancy of pieces of evidence the agent already has, behave like this. However, intuitively, most instances of reasoning with CMJs are not made as a result of exposure to this kind of evidence. Usually, when we reason with CMJs we do so on the basis of first order evidence, e.g. after being exposed to an argument that makes us change a specific CMJ we hold in a specific way. The result shows that the non-Bayesian must accept that out of any three

alternatives there are at least two alternatives such that the agent's CMJ regarding the moral superiority relation that holds between them is insensitive to such simple moral evidence.

I think it is clear that we want to allow for cases of reasoning that are based on this type of evidence. More generally, we want to allow for instances of reasoning that do not fall into either one of the two categories indicated above, and I think it is clear that anyone who is committed to the idea of reasoning with CMJs wants to allow for such cases too.

Notice that what made the transformation from impossibility to triviality so easy is the close connection, in the non-Bayesian framework, between experiences and their effects on agent's attitudes. For the non-Bayesian, granted that an experience is powerful enough to change one of the agent's CMJs, the change is binary. Thus, many different experiences (which constitute pieces of evidence with varying degrees of strength) that push in the same direction have exactly the same effect on the agent's attitudes. This is not the case for the Bayesian. For the Bayesian, stronger evidence leads to greater increase in degrees of belief.

So Bayesians use all the information to which they are exposed. Stronger evidence brings greater effect. Non-Bayesians, however, do not use all the information available to them. All pieces of evidence that do trigger a change bring the same change. Some information is lost. It is no wonder, then, that non-Bayesians get into trouble when it comes to reasoning.

Conclusion

The conclusion of this paper is quite strong: moral Bayesianism, the thesis according to which moral judgements are quantitative beliefs, is the only meta-ethical view that can account for moral reasoning. Nevertheless, the argument is sound.

Although the conclusion is strong, it should – I believe – be welcomed. As is demonstrated in section 1, it is not only very intuitive, but also carries with it many exciting and yet unexplored opportunities.

References

Bradley, R. List, C. (2009), Desire as Belief Revisited, Analysis, 69(1), 31-37.

Broome, J. (1991), Desire, Belief and Expectation, mind, 398, 265-257.

Broome, J. (2006), Reasoning with Preferences?, in Olsaretti, S. (ed.), Preferences and Well-being, Cambridge University Press.

Easwaran, K. and Fitelson, B. (2015), Accuracy, Coherence and Evidence, Oxford studies in Epistemology, 5, Oxford University Press.

Easwaran, K. (forthcoming), Dr. Trulove or: How I stop worrying and start loving Bayesian probabilities, Nous.

Field, H. (1978), A Note of Jeffrey Conditionalization, Philosophy of Science, 45, 3, 61- 367.

Greene, J. D. (2007). The secret joke of Kant's soul, in Moral Psychology, Vol. 3: The Neuroscience of Morality: Emotion, Disease, and Development, W. Sinnott-Armstrong, Ed., MIT Press, Cambridge, MA.

Hansson, S. O. (1995), Changes in preference, Theory and decisions, 38: 1-28.

Hansson, S. O. (2001), The Structure of Value and Norm, Cambridge University Press.

Jackson, F. and Smith, M. (2006), Absolutist Moral Theories and Uncertainty, Journal of Philosophy, 103, 267-283.

Jeffrey, C.R. (1992), Probability and the Art of Judgement, Cambridge Studies in Probability, Induction, and Decision Theory, Cambridge & New York: Cambridge University Press.

Joyce, M. J. (1998), A Nonpragmatic Vindication of Probabilism, Philosophy of Science, 65, 4, 575-603.

Kyburg, H.E. (1961) Probability and the Logic of Rational Belief, Middletown, CT: Wesleyan University Press.

Lepage, F. Morgan, C. (2011) Revision with Conditional Probability Functions: Two Impossibility Results, Dynamic Formal Epistemology, Springer Netherlands, 351, 161-172

List, C. (2012), The Theory of Judgement Aggregation: an Introductory Review, Synthese, 187 (1), 179-207.

Lockhart, T. (2000), Moral Uncertainty and its Consequences, Oxford University Press.

Mikhail, J. (2007), Universal moral grammar: theory, evidence and the future, Trends in Cognitive Science, 114, 143-152.

Nissan-Rosen, I. (2012), Doing the best one can: a new justification for the use of lotteries, Erasmus Journal for Philosophy and Economics, Volume 5, Issue 1, 45-72.

Oddie, G. (1994), Harmony, Purity, Truth", Mind, 103: 452-72.

Piller, C (2000), Doing What is Best, The Philosophical Quarterly, 50, 199.

Price, H. (1989), Defending Desire as Belief, Mind, XCVIII, 389, 119-127.

Sepielli, A. (2009), What to do when you do not know what to do?, Oxford studies in Metaethics, 4, Oxford University Press.

Singer, P. (1972), Famine, Affluence, and Morality, Philosophy and Public Affairs, 1:3, 229-43.

Singer, P. (2005), Intuitions, heuristics, and utilitarianism, Behavioral and Brain Sciences, 28, pp 560-1.

Smith, M. (2002). Evaluation, Uncertainty and Motivation, Ethical Theory and Moral Practice, 5, 305-320.

Sunstein, R. C. (2005), Moral Heuristics, Behavioural and Brain Sciences, 28, 531- 542.

van Roojen, M. (2009), Moral Cognitivism vs. Non-Cognitivism, Stanford Encyclopaedia of Philosophy.