

**IS CONVERGENCE IN DESIRES POSSIBLE?
A CRITICAL ASSESSMENT OF MICHAEL SMITH'S
MORAL REALISM**

**[A CONVERGÊNCIA EM DESEJOS É POSSÍVEL?
UMA AVALIAÇÃO CRÍTICA DO REALISMO MORAL DE
MICHAEL SMITH**

Lucas Mateus Dalsotto

Doutorando na Universidade Federal de Santa Maria
Bolsista CAPES

DOI: <http://dx.doi.org/10.21680/1983-2109.2018v25n47ID13254>

Natal, v. 25, n. 47
Maio-Ago. 2018, p. 143-161

Princípios
Revista de filosofia

E-ISSN: 1983-2109



Resumo: Neste artigo, pretendo avaliar se a abordagem de Michael Smith apoiando uma convergência substantiva entre os desejos idealizados dos agentes é convincente a ponto de nos mostrar que tal acordo seria possível. Smith afirma que a discussão moral tende a gerar acordo moral entre as opiniões dos agentes e que isso nos dá uma boa razão para acreditarmos que haverá uma convergência em desejos sob condições de racionalidade plena. Ele defende que a melhor explicação dessa tendência histórica é nossa convergência substantiva a respeito de um conjunto de verdades morais *a priori* pouco óbvias. Contudo, penso que ele falha em oferecer qualquer razão convincente para esperarmos que uma convergência substantiva seria alcançada. Argumento que o consenso de Smith sobre questões morais é menos provável do que o desacordo caso não haja nenhum padrão normativo de correção dos desejos que esperaríamos ter em condições de racionalidade plena. Se meu argumento é bem-sucedido, segue-se que a abordagem de Smith é um tipo teoria do erro de razões normativas, ao invés de um tipo de teoria realista.

Palavras-chave: Michael Smith; Realismo moral; Convergência; Desejos.

Abstract: In this paper, I intend to evaluate if Michael Smith's account in support of a substantive convergence among agents' idealized desires is compelling enough to show us that such agreement would be possible. Smith claims that moral argument tends to elicit moral agreement among agents' opinions and that this gives us a reason to believe that there would be a convergence in desires under conditions of full rationality. He maintains that the best explanation of that historical tendency is our substantive convergence upon a set of extremely unobvious *a priori* moral truths. However, I think he fails to provide us with any convincing reason to expect that a substantive convergence would be attained. I argue Smith's consensus on moral issues is less probable than disagreement if there is no normative standard for the correctness of the desires we would want ourselves to have in conditions of full rationality. If my argument succeeds it follows that Smith's account is a sort of error theory of normative reasons, rather than a realistic one.

Keywords: Michael Smith; Moral realism; Convergence; Desires.

Many authors have embraced in metaethics what may be called moral realism, that is, the view according to which moral claims purport to report facts and are true if they get the facts in the right way. Michael Smith (1994, p. 12) is one of those who have supported that idea. In *The Moral Problem*, he has held the thesis (which is widely known as *the objectivity thesis*) that our “moral statements of the form ‘It is right that I Φ ’ express a subject’s belief about an objective matter of fact, a fact about what is right for her to do.” He defends that the truth of this thesis rests on the idea that an agent has a reason to do a certain action just in case all fully rational agents would desire to do it in the same circumstances. So it is assumed by him that all fully rational agents would agree on a common set of idealized desires about what they have normative reason to do. But that is just a conceptual question. Smith also needs to show that a convergence in desires would emerge substantively. And in order to do that, he argues there is a historical tendency towards agreement. However, I suspect Smith fails to provide us with any compelling reason to expect that a substantive convergence would be attained. Keeping that in mind, in this paper I intend to make a critical assessment of his arguments in favour of the *presupposition* that there are objectively prescriptive features in the world. More specifically, I shall focus on the problem to know if his theory is able to account for how a substantive convergence among agents’ idealized desires would be possible.

Below, I shall begin with by considering Smith’s *rationalist thesis* and his *analysis of normative reasons* in support of the objectivity thesis¹ (section 1). After that, I shall expose his argument that a substantive convergence among agents’ idealized desires would emerge (section 2). Then I shall argue it fails to do this insofar as there is no decisive reason to persuade us that it

¹ Hereinafter, TOT (i.e., the objectivity thesis).

would be so (section 3). Lastly, I shall sum up the main ideas of this paper (section 4).

1. Rationalism and normative reasons

Smith's (1994, p. 185) core argument in favour of TOT rests on the view that if the analysis of moral rightness amounts to the idea of what we would want ourselves to do under certain ideal conditions of rationality, then our moral judgments are expressions of our beliefs about an objective matter of fact. And this is so because our moral claims are "expressions of our beliefs about what we have normative reason to do, where such reasons are in turn categorical requirements of rationality." However, it strikes me that such kind of defense of TOT is quite general and vague, so that it is hard to see how Smith explains it. I think it is necessary to look at his background arguments to have an accurate understanding of his account. In light of this, in what follows I shall dwell a little on them.

To show that TOT stands, Smith breaks his approach in a two-step argument: the rationalist thesis according to which the concept of moral requirement is the concept of a reason for action; and the analysis of normative reasons according to which what is right for us to do is what every rational agent would want herself to do.

In regards to the first step, the rationalist thesis defended by Smith (1994, p. 62) says that "if it is right for agents to do Φ in circumstances C , then there is a reason for those agents to do Φ in C ." A moral claim that it is right to do Φ in C can be true only if any rational agent has a reason to do Φ in C . To some extent, such thesis depends on the idea that, absent practical irrationality, agents will do what they judge they have reason to do. For it is a conceptual truth that if some actions are morally required for rational agents to do then we *expect* they will act accordingly. After all, this is what we can legitimately expect of rational agents as such. But this can be true just in case we imagine that moral

requirements are categorical requirements of rationality. Otherwise, it does not make sense. As Smith (1994, p. 90) claims, “[w]hat grounds the legitimacy of our expectations is the mere fact that people are rational agents. Being rational suffices to ground the expectation that people will do what they are morally required to do.”

According to Smith’s (1994, p. 5) argument, when we talk of reasons for action we quite generally take ourselves to be talking of a common subject matter. We assume that our moral reasons apply to others as well as ourselves, including our own possible and future selves. We believe agents facing the same circumstances would all have the same reasons for acting. For instance, when I say that doing Φ in C is right and you say it is not, we take it we disagree on the rightness of doing Φ in C . More than that, we take it that both of us cannot be correct, for it cannot be the case that every agent has a reason to do Φ in C and, at the same time, another reason not to do Φ in C . We do expect rational agents to make judgments truly and get the same conclusion about what they should do when they find themselves in the same circumstances. Therefore, if our moral judgments are claims about what the reason demands from us, then they are expressions of our beliefs about what we have reason to do. And this amounts to saying that “[o]ur concept of a moral requirement turns out to be the concept of a categorical requirement of rationality after all.” (Smith, 1994, p. 87).

Concerning the second step, it can be divided in two different parts. The first consists in stating that our claims about normative reasons are claims about “what we would desire to do if we were fully rational” (Smith, 1994, p. 136). At first sight, this idea may sound a little weird, because our normative reasons are beliefs about what we would desire to do in optimal conditions of deliberation. But it is not. On the one hand, having a normative reason is a matter of believing in a sense that valuing something is a matter of believing in something. On the other, the right content

of a belief is to be understood in terms of desirability, which is to say, what we would want ourselves to do if we were fully rational.

To make this clear, Smith (1994, p. 150-151) asks us how we should decide what to do in cases in which we are hesitant about what to do. For him, the best answer for such question should be 'to ask for advice'. Yet it is not any sort of advice. It should be an advice provided by one better situated than ourselves in a given set of circumstances, and one who knows us well. On his view, these two conditions can be met through the consideration that our idealized counterparts are the best people to give us advice about what it is required us to do. What we have normative reason to do in circumstances C is what "we, not as we actually are, but as we would be in a possible world² in which we are fully rational [...], would want ourselves to do in those circumstances."

In that sense, it is right for agents to do Φ in the world W just in case their fully rational counterparts in an idealized world W^* would desire to do it as well. The analysis of the claim that an agent has a normative reason for doing Φ in C is not a claim about her *actual* desires, but instead one about her *idealized* desires. The truth of the proposition 'John has a reason for doing Φ in C ' does not imply that John has some personal desire which will lead him to do Φ in C . "What it implies is rather that he *would* have some such a desire if he were fully rational" (Smith, 1994, p. 165). John's reason for doing Φ in C is given by the content of the desires he would have if he were in optimal conditions for deliberating.

The second part of the second step is constituted by the idea that our claims about normative reasons are *non-relative* to the agent. When two different people are talking about reasons for acting, they presuppose they are talking about the same thing. In a way, the reason that Smith (1994, p 168) offers in support of a

² For an appropriate overview about that, see the entry on "Possible Worlds" in *Stanford Encyclopedia of Philosophy* (Menzel, 2016).

non-relative concept of normative reasons is that other normative concepts such as *truth*, *support*, and *entailment* do not give rise to claims with relative truth conditions. If I say ‘*a* implies *b*’ and you say it does not, we suppose that one of us is wrong because we understand the concept of entailment in a non-relative sense. And this strongly reinforces the thesis that “propositions have normative force *simpliciter*, not just normative-force-relative-to-this-individual or relative-to-that.” According to Smith (1994), it follows that the truth of a normative reason claim presupposes that, under conditions of full rationality, we would all get the same conclusion about what is to be done in the various situations we might face. There would be a convergence in desires among all fully rational agents.

However, a point needs to be made here. Even though Smith’s (1994, p. 168) concept of normative reasons is non-relative, sometimes we find ourselves in situations in which we say something like ‘That can be desirable for you, but it is not desirable for me’. In that scenario, one might wonder: ‘How does Smith understand this sort of *relativity*?’ Well, he argues that this sort of ‘relativity’ does not undermine the idea that some of our moral claims are categorical. On his view, our choices and preferences may sometimes be a relevant feature of our circumstances, but it depends on whether fully rational agents would desire to do the same thing if they were in the same circumstances and in the same position we find ourselves. Even if reasons are non-relative in the crucial sense at issue, among the variables that can rationally justify our choices are considerations that are *agent-relative* and considerations that are *agent-neutral*, to use Derek Parfit’s (1984, p. 26-27) terms.

Let us consider Smith’s (1994, p. 169) example by way of illustration. Imagine you are standing on a beach and two people are drowning to your right, and one is drowning to your left. Faced with this scene, either you can swim right and save two, in which case the one on the left will drown, or you can swim left and save

just one, in which case the two on the right will drown. Imagine also that the one who is drowning to your left is your child. In light of these available variables, it seems clear that you would choose to swim left and save one. You would possibly justify what you did by claiming that ‘the one on the left was my child, while the two on the right were completely strangers to me’. So what would my reaction be concerning your situation?

From one perspective, I would say ‘this is a reason for you, but it is not for me’. Since the three people are strangers to me, I would probably choose to swim right and save two people instead of one. But from another, what is a reason for you can indeed be a reason for me. If I had been in your place and the one on the left had been my child, then surely I would have been able to justify my choice using the same reason you had taken to justify yours. It seems very plausible to assume that in this case we have both sorts of considerations. On the one hand, you have an agent-relative consideration in the sense that in those circumstances you had a personal reason to save your child. On the other, you have an agent-neutral consideration in the sense that in those circumstances anyone would have a normative reason to save her own child (Smith, 1994, p 170).

Anyway, what it is worth noting with respect to the idea that our normative reasons are non-relative is that to say the features of liberty and justice, for instance, merit our love and devotion implies to give our reasons for thinking this is so. Those features deserve our love and devotion only if the reasons we provide in their support have appeal to any creature capable of asking the question ‘Should I care about liberty and justice?’, not just to ourselves (Smith, 1989, p. 102). Moral beliefs are beliefs about some non-relative matter of fact, and the search for reasons in support of our moral beliefs is the search for reasons that would convince every rational agent to assume such beliefs.

Having reached at this point of Smith’s argumentation, it follows from the conjunction of the rationalist thesis and the ana-

lysis of normative reasons that if it is right to do Φ in C and if x is a fully rational agent, then every agent x would desire to do Φ in C . On Smith's view, that is enough to establish the truth of TOT and show that our moral judgments suppose the existence of categorical requirements of rationality. Nevertheless, now he needs to provide us with a compelling reason to believe in the idea that rational agents would converge in desires and so our moral claims are not based on a massive error of presupposition. In the next section, I shall expose Smith's reasons in favour of such idea.

2. A tendency towards agreement

As just seen, even though Smith's concept of normative reasons is non-relative, the source of what we have reason to do is given by the content of the desires we would have if we were fully rational. This amounts to stating that the plausibility of Smith's realist picture of moral truth requires and presupposes that all agents would substantively converge on the same set of desires under conditions of full rationality. Agents should have the same desires about what needs to be done and desired in any morally relevant situation of their lives. In case such convergence in desires is not possible our moral talk cannot be true and legitimate, since it implies systematic error.

The possibility of such convergence thereby depends on the agents' engagement in the process of systematic justification of their desires in order to achieve a common set of reasons that they can share. All of them need to evaluate whether or not their actions – of course, actions that matter to the moral domain – fulfill the requirement that the right thing to do is what they would desire to do under conditions of full rationality. Yet it is worth remarking that the convergence required at the level of desires is not about how each agent structures her own life in her own world. Agents undoubtedly will find themselves in different circumstances from each other, circumstances that may be constrained by their behaviors, personal skills, and attachments in

their respective worlds. The convergence required rather concerns only the idealized desires that every fully rational agent would have as a normative reason for acting (Smith, 1994, p. 173).

In defense of this idea, Smith (1994, p. 187) is quite economical, offering little discussion on this important point of his account. Anyway, his core argument is that the “empirical fact that moral argument tends to elicit the agreement of our fellows gives us a reason to believe that there will be a convergence in our desires under conditions of full rationality.” He holds that the best explanation of that historical tendency is our substantive convergence upon a set of extremely unobvious *a priori* moral truths. The tenability of these unobvious *a priori* moral truths requires a convergence about what fully rational agents would desire to do in certain ideal circumstances. If moral facts are facts about the agents’ normative reasons and those reasons are constituted by their idealized desires, then the supposition that some of our moral claims are objectively correct cannot succeed unless we substantively converge in desires.

Moreover, Smith (1994, p. 188-189) also provides other three supplementary empirical reasons in favour of the tendency towards agreement. The first is that there are large areas of agreement even alongside areas of moral disagreement. Sometimes by focusing only on moral disagreement we ignore the substantial agreement we have already achieved, and the fact we share *thick* evaluative language concepts such as courage, brutality, and loyalty (see, e.g., Williams, 1985, p. 129). And the prevalence of such concepts shows us that there is a considerable agreement among people about what is right and wrong to do. The second reason is that, although current moral disputes appear sometimes deadlocked, we need to remember that in the past similarly entrenched disagreements were solved, among other things, through a process of moral argument. According to Smith (1994, p. 188), “we must not forget that there has been considerable moral progress, and that what moral progress consists in is

the removal of entrenched disagreements of just kind that we currently face.” And the third reason is that sometimes entrenched disagreements can be described in ways that make it look less entrenched. Some moral disagreements can be explained as clearly arising from lack of free and rational debate.

In sum, Smith’s (1994, p. 189) argument is that we have good historical evidences to believe in the “possibility of an agreement about what is right and wrong being reached under more idealized conditions of reflection and discussion.” And there is thus no relevant reason to believe that our moral talk presupposes wrongly the idea that the world contains objectively prescriptive features, as Mackie (1977, p. 27-30) advocates.

3. Is convergence in desires possible?

Even though convergence is widely assumed as possible, whether it will occur is always *sub judice*. But the success of Smith’s realist picture of moral truth depends on such convergence be possible. His position faces the following dilemma: (i) either the desires of all fully rational agents converge on a common set of desires (ii) or our claims about normative reasons are always false and so an error theory is the way to go. Although Smith (1994, p. 189) remains open-minded about (ii), he is clearly optimistic about (i). Yet I am not. If those reasons exposed in section 2 are all what Smith has to say in support of the agreement among agents’ idealized desires, then I think there is no decisive reason to expect that one would emerge by means of a free and rational debate. In what follows, I shall argue Smith’s consensus on moral issues is less probable than disagreement if there is no normative standard for the correctness of the desires we would want ourselves to have in conditions of full rationality.

I believe that one striking objection to Smith’s historical and inductive argument for agreement is that made by David Sobel (1999). On my view, Sobel’s (1999, p 146) objection makes it clear that the reasons offered by Smith in favour of a substantive

convergence in desires are not persuasive enough to make us believe that we would get it in the future. A relevant part of the moral consensus we actually have is the result of factors other than the free and rational debate. Some of our historical moral discussions that have produced agreement have been “(1) factually and logically imperfect, (2) addressed to those poorly positioned to object [and] (3) to those who share substantial common moral vocabulary, moral education, and cultural identification, and (4) offered by those who are persuasive for reasons other than the cogency of their position.” Moreover, any agreement reached through these four routes does not constitute reliable evidence that fully rational agents would achieve the same set of idealized desires³ after due process of deliberation.

Another point that Sobel (1999, p. 146) calls us attention to concerns Smith’s idea that the agreement to be reached is not simply among all *real* agents’ desires, but also among all *possible* agents’ desires. This implies to say Smith needs to show us that a substantive convergence in desires is possible no matter how different are the real and possible agents’ vocabulary and moral education. More than that, he needs to show us that it would happen in those circumstances for the right reasons, where this means via a free and rational debate. As said above, the agreement cannot be the consequence of cultural hegemony or religious authority, for instance. Because of this, Sobel (1999, p. 147) holds “Smith needs to show that history provides a good inductive case that all fully rational agents, no matter how initially divergent they are prior to becoming fully rational, will agree on moral matters.” Yet, as we can already imagine, Sobel is not persuaded by Smith’s argument.

In light of these considerations, I understand that in order to show that moral disagreements can disappear over time Smith

³ Obviously, Sobel (1999, p. 146) is not denying that those four causes play an important role in the production of agreement on moral matters. He is just saying that they will not constitute “inductive evidence for Smith’s case.”

needs to provide us with a non-empirical reason that a convergence in desires would be possible. As presented earlier, he is trying to convince us that there is a historical tendency towards agreement and that this tendency is inductive evidence in favour of convergence. However, the issue is that it is needed to show that there is a necessary convergence among agents' normative reasons through a rational argument despite disagreement, since a consensus on moral matters can be explained by means of several non-rational processes, as Sobel (1999, p. 146) points out. And at least as built in *The Moral Problem*, Smith's (1994) theory does not seem to have such non-empirical reason for doing such a thing.

Hence, my argument against Smith's account turns out to be that convergence is less probable than disagreement if there is no normative standard for the correctness of the desires we would want ourselves to have if we were fully rational. Keeping in mind Sobel's remarks, it seems reasonable to state that agents start off from different desiderata in their deliberations, since sometimes they may have radically divergent vocabularies and moral educations. And by virtue of their psychological inclinations⁴, agents may form desires about what they would want themselves to do in conditions of full rationality differently. The fact agents idealize some desires does not mean that the outcome is the same for all of them, for they may react differently on the same facts. Notice I am not referring to the divergence among agents' preferences for, say, wine or beer, but rather to the possible divergence among agents' desires that make true those moral claims that are in fact true. Yet one might ask: 'What do I mean by normative standard for the correctness of desires?'

By normative standard for the correctness of desires I mean the idea that there should be some rational criterion to aid in the free and rational debate of which desires agents have a normative reason to want in conditions of full rationality. I think if there is no

⁴ By *inclination* I mean things like *capacities, personal skills, drives* and so on.

'point' in which agents can agree prior they engage in a rational debate, then the probability that all ideal agents would end up with the same set of desires after due process of deliberation is extremely low. But if we have a rational criterion, say, *R* that helps agents in the process of systematic justification of their desires, then the probability that all ideal agents would converge is high. In case agents arrive at a different set of idealized desires, this can probably be explained by the misuse of *R*, which offers a pattern of correctness of which desires agents would want themselves to have if they were fully rational. At least as I see the question, the defense of this idea is completely consistent with Smith's theoretical framework as a whole.

Of course, an important issue here is to determine what *R* might be. Looking back at the philosophical literature, it might be a standard for maximizing utility or following the categorical imperative, for example. To show how this view could improve Smith's account, let us imagine he incorporates into his argument of the tendency towards agreement the idea that all ideal agents would converge in desires because this is in line with the normative standard of following the categorical imperative. In that case, fully rational agents would agree on a common set of desires not because they share thick evaluative language concepts, but because it is rational for them to follow the categorical imperative. A consensus among agents' idealized desires would be attained because there is a normative standard for the correctness of desires. From that outlook, Smith would be able to provide us with a stronger reason than those presented in section 2 to believe that a substantive convergence would emerge. Hence, I have no doubt Smith is better off if he can show that agreement is something rationality required, not some historical propensity that makes us converge on certain moral matters.

A good way of illustrating how Smith's idea of agreement is too much to believe without a normative standard for the correctness of desires is considering David Enoch's (2007, p. 105-106) thought

experiment. Suppose I ask people to randomly choose a whole number between, say, 0 and 10. Suppose further that all of them actually chose the same number under conditions of full rationality. What would we think about that?

Wouldn't this be amazing? Given that there are infinitely many options, wouldn't such convergence cry out for an explanation? Without such explanation, wouldn't convergence be utterly miraculous, and so utterly incredible? At the beginning of my experiment, before the results are in, would you be willing to bet money on the emergence of such an amazing convergence? Well, perhaps as the thought experiment stands such convergence would not be *all that* amazing. Perhaps, for instance, many will be drawn to the rather simple and symmetric [5]. Or perhaps something in our human hard-wiring makes the answer [7] come naturally to us, or something of this sort. But suppose I conduct the experiment not just among all persons, but rather among all *possible* persons.

Enoch (2007, p. 106) suggests Smith might account for this miracle by claiming that there are “desire-independent facts about what is desirable” to do. If this were the case then Smith might state that, say, 7 is an appropriate response because it is a perfect number. Many cultures and religions see 7 as the representation of perfection and completeness. For instance, there are seven gifts of the Holy Spirit, seven chakras, seven rainbow colors, seven musical notes, and so on. In that scenario, it would not be too surprising if all possible people converged on 7. Probably, the reason for this would be that 7 is a number which is seen in many cultures and religions as the number of perfection and completeness. All possible people would be able to choose 7 because it is desire-independent.

Nevertheless, as must be clear at this point, this route is not open to Smith, for he does not believe in desire-independent facts about what is desirable to do. On his view, facts about the desirability of acting in a certain way in the world are constituted by facts about the desires we would ideally have in relation to

something in it. As seen earlier, Smith (1994, p. 152) understands that “facts about what is desirable for us to do are constituted by facts about what we would advise ourselves to do if we were perfectly placed to give ourselves advice.” What constitutes the fact that something is a normative reason for one is a relation between that thing and one’s idealized desires. Yet if there are no desire-independent facts about what is desirable to do, then Smith’s theory seems to be unable to account for how such miracle is possible.

Unlike Enoch’s (2007, p. 106) thought, I do not think that in order to show that a substantive convergence in desires is possible Smith should maintain that there are desire-independent facts about what is desirable to do. This would lead him to give up much of his theoretical framework and this is not necessary. But I do think Smith’s theory needs to offer us a normative standard for the correctness of the desires we would want ourselves to have if we were fully rational. Without such standard, just as a consensus among people’s bets would be miraculous, so a consensus on the same set of desires among all fully rational agents would be too. And pending other reasons for believing in the substantive convergence, there would be probably none.

Anyway, the point that matters now is to find out what happens to Smith’s theory in case it cannot give any decisive reason to persuade us that an agreement among agents’ idealized desires would be possible. Sometimes Smith (1994, p. 189) even recognizes that if there is no consensus on certain moral facts, “then we might well quite justifiably come to think that Mackie⁵ was right after all.” And this makes sense because by failing to

⁵ Mackie’s (1977, p. 37-38) argument is that moral discourse is committed to the idea that moral facts refer to properties, relations, and moral entities, but that such properties, relations, and moral entities, in fact, do not exist in the world. He states that “if there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe.”

show that an agreement among agents' idealized desires would emerge, Smith fails to show that our claims about normative reasons can be objectively correct. And by doing so, his account collapses into the idea that our moral judgments are always false, given they are based on a massive error of presupposition, that is, the presupposition the world contains objectively prescriptive features. I therefore endorse the conclusion advocated by Richard Joyce (2001, p. 86) that Smith's theory "will amount to an error theory of normative reasons."⁶

In brief, in this section I tried to show that Smith's account seems to have failed at least on the point of offering a reason in support of the idea of a convergence in desires. His historical and inductive argument is not convincing enough to make me think that a consensus on moral issues would be attained via a free and rational debate. I also tried to argue that Smith's theory is better off if he incorporates into it the view that there is a normative standard for the correctness of the desires we would want ourselves to have if we were fully rational.

4. Concluding remarks

The aim of this paper has been to find out if Smith's theory is able to account for how a substantive convergence among agents' idealized desires would be possible. In *The Moral Problem*, Smith (1994, p. 187) maintains that a consensus on moral matters would emerge because there is a tendency towards agreement and the best way to explain it is from "our convergence upon a set of extremely unobvious *a priori* moral truths." However, I have argued that his historical and inductive argument in favour of convergence is not compelling enough to make me believe it would be possible. I have also claimed that Smith is better off if he incorporates into his theory the idea that there is a normative

⁶ Of course, the reasons by which I come to this conclusion are quite different from those he comes to.

standard for the correctness of the desires we would want ourselves to have if we were fully rational. As a result, I have endorsed Joyce's (2001, p. 86) conclusion that "all Smith's labors will be for the error theorist in the end."

Acknowledgements

I would like to thank Sergio Tenenbaum (University of Toronto), João Carlos Brum Torres (UFRGS/UCS), Ricardo Bins di Napoli (UFSM), and Flávio Williges (UFSM) for valuable remarks and comments on drafts of this paper. A previous version of it was presented at the *XVII Encontro da Associação Nacional de Pós-Graduação em Filosofia* (ANPOF). Then I am also indebted to the participants of the GT Ética, especially Alcino Bonella (UFU), Rogério Picolli (UFSJ), Vitor Somavilla de Souza Barros (UFMG), and Adelino Ferreira (UFMG). Finally, I would like to express my gratitude to CAPES for the financial support.

References

- ENOCH, David. Rationality, coherence, convergence: a critical comment on Michael Smith's "Ethics and the A Priori". *Philosophical Books*. v. 48, n. 2, 2007, p. 99-108.
- JOYCE, Richard. *The myth of morality*. Cambridge: Cambridge University Press, 2001.
- MACKIE, John. *Ethics: inventing wrong and right*. Harmondsworth: Penguin, 1977.
- MENZEL, Christopher. Possible worlds. In: ZALTA, Edward N. (Ed.). *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition). Available at: <<https://plato.stanford.edu/archives/win2016/entries/possible-worlds/>>. Access: 8 dec. 2017.

PARFIT, Derek. *Reasons and persons*. New York: Oxford University Press, 1984.

SMITH, Michael. Dispositional theories of value. In: THE ARISTOTELIAN SOCIETY. *Proceedings of Aristotelian Society, Supplementary Volumes* – v. 63. Oxford: Oxford University Press, 1989. p. 89-111.

SMITH, Michael. *The moral problem*. Oxford: Wiley-Blackwell, 1994.

SOBEL, David. Do the desires of rational agents converge? *Analysis*. v. 59, n. 3, 1999, p. 137-147.

WILLIAMS, Bernard. *Ethics and the limits of Philosophy*. Cambridge: Harvard University Press, 1985.

Artigo recebido em 5/12/2017, aprovado em 27/02/2018