

MORAL OBJECTIVITY, CONVERGENCE, AND DESIRES: TWO (POSSIBLE) OBJECTIONS TO MICHAEL SMITH'S CONSTITUTIVISM

Lucas Dalsotto

Universidade de Caxias do Sul

Resumo: A fim de explicar como é possível haver objetividade moral, Michael Smith tem sugerido que o moralmente correto deve ser compreendido em termos de razões para ação, onde isto significa o que um agente plenamente racional e coerente desejaria idealmente fazer. Para explicar esta ideia, ele tem defendido uma abordagem constitutivista segundo a qual existiriam certos desejos que são constitutivos da agência ideal, tais como os desejos de ajudar a garantir e de não intervir no exercício da capacidade deliberativa de conhecer o mundo em que vivemos e da capacidade deliberativa de alcançar nossos desejos nele. Do argumento de Smith segue-se então que a convergência entre os desejos dos agentes plenamente racionais e coerentes emergirá no decorrer do tempo e espaço, garantindo assim a objetividade moral. Mas suspeito que, ao menos do modo como a discussão está atualmente estabelecida, a abordagem constitutivista de Smith está em apuros. Para mostrar isso, sugeri que sua teoria é vítima de dois problemas em particular, e então argumentarei que ela falha em explicar adequadamente a natureza objetiva da moralidade.

Palavras-chave: Smith, constitutivismo, objetividade, convergência.

Abstract: In order to explain how moral objectivity is possible, Michael Smith has suggested that moral rightness should be understood in terms of reasons for action, where this means what a fully rational and coherent agent would ideally desire to do. To account for that idea, he has defended a constitutivist approach according to which there might be certain desires that are constitutive of the ideal agency, such as the desire to help to ensure and not to interfere with the exercise of the deliberative capacity to know the world in which we live and the deliberative capacity to achieve our final desires in it. It then follows from Smith's argument that convergence in fully rational and coherent agents' desires will emerge over time and space, ensuring thus moral objectivity. Yet I suspect that, at least according to the current state of the art, Smith's constitutivism approach is in trouble. To show that, I shall suggest that his theory is victim of two problems in particular, and then argue it fails to explain adequately the objective nature of morality.

Keywords: Smith, constitutivism, objectivity, convergence.

Introduction

One of today's most interesting debates in metaethics refers to the understanding of the objective nature of morality. In a general sense, the phenomenology of our moral experience leads us to believe that if two agents are in the same conditions and act in the same way, either both acted properly or improperly. For example, if I assert 'stealing is wrong in the circumstances

C and you assert ‘stealing is right in the circumstances C ’, then we believe that one of us is making a misjudgment. The idea behind this reasoning is that, despite any material difference in the conditions of a particular fact, the reasons that apply to a person as an appropriate response to this fact must be enforced as an appropriate response to any other person. When we claim from someone that ‘stealing is wrong in the circumstances C ’ we are assuming that she can recognize and share the reasons we have to hold that ‘stealing is wrong in the circumstances C ’.

In the recent philosophical literature, Michael Smith (1994) has suggested we should understand the objective nature of morality in terms of what a fully rational and coherent agent would desire that her real-world counterpart does. He presupposes it is a *platitude* about normative reasons that what is desirable that we do is what we would desire to do if we are fully rational and coherent (SMITH, 1994, p. 165). Agents who face the same circumstances all will have the same reason for acting. There would be a convergence at the level of hypothetical desires about what is to be done in the various circumstances in which rational agents could find themselves. So the question I want to answer in this article is the following: Does Smith’s theory account for how agents would ideally converge on the same set of desires over time and space?

In order to respond that question, I shall begin by presenting the general landscape in which Smith’s theory is attached to (section 2). Then I shall expose how his constitutivist approach is thought to explain the objective nature of morality (section 3). After that, I shall argue that his constitutivist approach is victim of two objections in particular (section 4). And finally, I shall sum up the main ideas developed in this article (section 5).

1. The Background Discussion

In his book *The Moral Problem*, Smith (1994, p. 12) maintains we should understand the objective nature of morality through the following proposition:

- (1) Moral judgments of the form ‘It is right that I do ϕ ’ express a subject’s belief about an objective matter of fact, a fact about what is right for her to do.

In a very schematic way¹, Smith’s argument to explain (1) is based on the view that it is right for agents to do ϕ in the circumstances C if and only if

¹ I will not discuss Smith’s background ideas in further detail.

there is a normative reason for them to do ϕ in the circumstances C . That view is linked to two basic ideas: an analysis of moral rightness in terms of reasons for action; and a conception of moral rightness according to which it is right for an agent to do something just in case she has a normative reason to do it. From that outlook, moral rightness should be understood in terms of facts about normative reasons that people have, which is to say, what would be good and desirable to do if they were fully rational and coherent² and where these desires are of an *appropriate substantive kind*³. Thus, believing that an act is desirable “is believing that one’s ideally rational fully informed and coherent self would desire that one imperfect as one actually is perform it.” (ENOCH, 2007, p.99).

To make this more clearly, it is important to notice that the analysis of moral rightness offered by Smith is founded on a relation among an action-type, a person, and a kind of circumstance. Let us suppose that an action-type is ‘telling the truth’, a person is ‘John’, and a kind of circumstances is ‘ C ’. In that sense, we can derive from that idea the proposition that ‘John has a normative reason to tell the truth in the circumstances C if, being fully rational and coherent, he would desire to tell the truth in C and if ‘telling the truth in C is an act of an appropriate substantive kind. The general proposition that follows from this is that presumably ‘any person, if fully rational and coherent, would desire to tell the truth in the circumstances C ’.

In a sense, what Smith is trying to do is to defend that moral objectivity is possible without resorting to any metaphysically queer entities. He believes that if there is any normative reason at all, then agents who face the same circumstances all will have the same reason for acting. There would be a convergence at the level of hypothetical desires about what is to be done in the various circumstances in which rational agents could find themselves. Everyone could have the same set of idealized desires (on fundamental moral issues) if they engage in the process of systematic justification of their desires. There would be a convergence on norms and principles of practical reason, and then the requirements of rationality about what we have normative reasons to do will be requirements of morality as well (COPP, 1997, p. 45).

² Smith (1994, p. 156) follows Williams’ (1981, p. 102-3) conception of full rationality. According to such conception, a fully rational and coherent agent is someone who (i) has no false beliefs; (ii) has all the relevant true beliefs; (iii) tries to achieve a justification system to ensure consistency between beliefs and desires; and (iv) deliberates properly using the variables of each situation

³ By *appropriate substantive kind* Smith (1994, p. 40) means, for example, that moral reasons are concerned with the promotion of human welfare (see, e.g., FOOT, 1972, p. 313), social justice (see, e.g., RAWLS, 1971, p. 3-4), equal respect among people (see, e.g., DWORKIN, 1977, p. 179-83; KIMLICKA, 1989, p. 21-9) or the like.

But the problem that arises now is the following: How is such convergence substantively possible? David Enoch (2007), for instance, is skeptical about Smith's substantive convergence among agents' idealized desires. He argues that if there is no point in which agents can converge prior they engage in a rational argument, then the probability that all agents will end up with the same set of desires after a deliberation process is extremely low⁴. Roughly speaking, Enoch's (2007, p. 105-6) argument is made by means of an experimental thought. He asks us to imagine a situation in which I ask people to randomly pick up a rational number between, say, 0 and 1.

And [imagine further] all of them actually come up with the same number. 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 – or indeed the fate of your favorite theory – on the emergency of such an amazing convergence? [...] Perhaps, for instance, many will be drawn to the rather simple and symmetric 0,5. Or perhaps something in our human hard-wiring makes the answer 0,7 comes naturally to us, or something of this sort. But now suppose I conduct the experiment not just among all persons, but rather among all *possible* persons. People in different possible worlds differ in their hard-wiring, and indeed in their attraction to symmetry. Absent some explanation, a convergence of all *possible* persons would thus clearly be too much to believe.

To some extent, the aim of the next section will be presenting Smith's reply to this kind of skepticism about how agents' ideal counterparts would substantively converge on the same set of desires.

2. The Constitutivist Approach

Smith (2011; 2013; 2015) has recently published a couple of papers in order to develop a novel form of *constitutivism*⁵. He believes there could be a convergence among agents' desires the extent to which it is partially constitutive of being fully rational and coherent having certain dominant final desires, which are fixed by the internal norms to the concept of agency. All

⁴ By a similar statement, see Joyce (2001, p. 85-96).

⁵ A focal point in recent work on practical reason is the idea that we might ground normative claims in facts about the nature of action and agency. This position is widely known as *constitutivism*. The core argument of this view is that certain normative claims apply to us merely because we are agents, because universal norms of assessment are grounded in nothing more than the nature of agency itself. To make this more clearly, imagine a competitive soccer match. There are many rules and regulations, but to engage fully in the activity of playing competitive soccer, one must attempt to score more goals than one's opponent. This *aim* – or *principle* – seems to constitute the practice of soccer match. Without it, one is simply not in the business of playing soccer. Such analogy appears to illuminate the way in which a constitutive *aim* – or *principle* – guides agents' actions and provides a standard of success and assessment. For more about that topic, see Millgram (2001, p. 5-15).

ideal agents would have the same dominant final desires because having them is part of having an ideal psychology.

With that in mind, in this section I intend to introduce the main arguments of Smith's constitutivism to explain the objective nature of morality and so prove that (1) is true⁶. But before doing that, I would like to draw attention to three aspects of his proposal. First, the constitutivist approach is an attempt to provide a reason to believe in the substantive convergence among agents' idealized desires and thus deny some skeptical conclusions like Enoch's one about the realist character of his theory⁷. For that reason, I will read the constitutivist approach as a natural extension of his previous framework. Second, unlike constitutivists such as David Velleman (2009, p. 179-86) and Christine Korsgaard (2008, ch. 2), Smith (1994, p. 185-6) defines reasons not directly in terms of what is constitutive of the agency, but instead in terms of what is constitutive of a fully rational and coherent agency⁸. And third, Smith's constitutivism seems to be a work in progress⁹, so my presentation and analysis are restricted to the current state of the art.

To begin, Smith's constitutivism rests on the following thesis.

(2) Moral requirements can be reduced to rational requirements.

Smith's aim is to give a rational foundation for morality by providing an account of rational requirements that makes it plausible that moral requirements reduce to rational requirements. As reasons for action have recognizably moral content, they are moral reasons about what we are obligated to do. In order to better understand Smith's constitutivism, I will follow Bukoski's (2016, p. 117) suggestion and then divide it into three main parts. The first part consists in introducing a given conception of action and ideal agency. The second consists in exposing the normative significance of the ideal agency. And the third part consists in deriving the various subsidiary desires that are constitutive of the ideal agency. Let me now say more about each of them.

With regard to the first part, Smith's account is founded in the causal theory of action according to which what makes an agent's movement of her

⁶ My presentation is based mostly on his paper "A Constitutivism Theory of Reasons: Its Promise and Parts".

⁷ See Smith (2015, p. 18).

⁸ Constitutivists such as Velleman and Korsgaard usually try to show that some kind of norm – e.g., the Categorical Imperative, in Korsgaard's case – is implicit in being an agent, and then she needs to follow it, on pain of not being one. From that perspective, agency is inescapable (see, e.g., FERRERO, 2009, p. 308) in a sense we cannot but be agents, that it is not optional.

⁹ According to Smith (2013, p. 28), "[m]uch work still needs to be done in filling out the details, of course. But hopefully enough has been said to make it clear how well-placed Constitutivism is to deliver on its promise."

body an action is the fact that this movement is produced in the right way by two psychological states, namely: some final desire and some means-end belief to realize that final desire. This conception of action, in turn, supports a specific conception of ideal agency, which states that what is constitutive of the agency is the capacity for action. Put in slightly more specific terms, Smith (2013, p. 18) believes that “someone is an agent in virtue of being capable of action, which is to say, by virtue of having the capacity to realize her final desires, given her beliefs.”

He defines the notion of ideal agency according to Judith J. Thompson’s (2008, p. 21-2) *goodness-fixing kinds*, where goodness-fixing kinds are kinds whose nature fixes a standard of success for things of that kind. To clarify such an idea, Thompson gives us as examples the toaster and the burglars. Toasters are devices for warming and browning bread so that you can enjoy eating it. In that way, the standard of success for a toaster is doing all this without burning the toast, making it much more enjoyable to eat. Burglars, in turn, are people who make their living by stealing things. So the standard of success for a burglar is doing this without getting caught.

Analogously, Smith (2013, p.10) believes that a good agent is who uses the two constitutive capacities of the agency, namely: (i) the capacity to form rational beliefs about the world, which I call the *knowledge acquisition* capacity, and (ii) the capacity for means-end coherence in light of one’s beliefs about how to achieve one’s final desires, which I call the *instrumental reasoning* capacity. It, therefore, follows that the standard of success for a good agent is exercising and using the deliberative capacity to know the world in which she lives and the deliberative capacity to achieve her final desires in it.

Smith (2013, p. 19) then seeks to derive some agent’s dominant final desires through his view about ideal agency. He states that ideal agents will have a maximally coherent and unified psychology, but the exercise of the *knowledge acquisition* and *instrumental reasoning* capacities can sometimes undermine coherence. They can direct us in opposite directions in a wide range of circumstances. Let us imagine that an agent’s ideal counterpart desires that she at this moment believes that ϕ . On the one hand, the exercise of the *instrumental reasoning* capacity will lead her to take the means to achieve ϕ , where this means to say she believes that ϕ whether or not ϕ is true. On the other hand, the exercise of the *knowledge acquisition* capacity will lead her to know the world in which she lives, where this means to say she believes that ϕ if and only if ϕ is true.

If this makes sense, then the use of these two deliberative capacities can produce a contradiction into the ideal agent’s psychology in a sense that

one capacity does not cohere with the exercise of the other. Nonetheless, Smith holds that all ideal agents have a dominant desire *not to interfere* in the exercise of their *knowledge acquisition* capacity, where a dominant desire is to be read as one that overrides other desires. The dominant desire not to interfere in the exercise of their *knowledge acquisition* capacity would remove all potential conflict between the exercise of agents' deliberative capacity to know the world in which they live and the exercise of their deliberative capacity to achieve their desires in it. This dominant desire privileges the *knowledge acquisition* capacity insofar as a *final* desire is a desire for something for its own sake, whereas an *instrumental* desire is an object of which is desired as means to achieving a *final* desire. Thus, the dual deliberative capacity of the ideal agency also induces more coherence and unity in an agent's psychology through the dominant final desire not to interfere.

Smith (2013, p. 23) also derives a second agent's dominant final desire through his view about ideal agency. The point is that there may be a deep divergence between an agent's relationship to her own current beliefs and desires, and those she will have in the future by virtue of the fact that her present beliefs and desires can undermine her future self's exercise of her two deliberative capacities. Imagine a situation in which I know I will have a debilitating headache, and now I have available a pill which, if taken later, would remove my headache in the future. Imagine also that I do not like to take medicine, and therefore I decide not to take the pill later. In such a case, wouldn't I be undermining the exercise of my two deliberative capacities in the future? (BUKOSKI, 2016, p. 133).

According to Smith's (2013, p. 24) argument, I would. He maintains that an agent would have to desire that he now does what he can *to help to ensure* that his future self has the required capacities so that she can be an ideal agent. "[A]n agent's being ideal at a time requires her to be, at that time, such as she needs to be in order to be ideal, not just at that time, but also at later times." In the example earlier, I would have to desire to take the pill so that my future self would take it, on pain of not being an ideal agent at all. Of course, an ideal agent cannot guarantee at this moment that she will be ideal later, given that this depends on certain future circumstances of the world and on how she will use her two deliberative capacities. In any case, Smith believes that all ideal agents have a dominant desire to help to ensure the future exercise of their *knowledge acquisition* and *instrumental reasoning* capacities.

Then the upshot of this view is that each ideal agent has the following dominant final desires: (a) a dominant desire not to interfere with the current and future exercise of the deliberative capacity to have knowledge of the world in which she lives, and (b) a dominant desire to help to ensure in the future the

deliberative capacity to have knowledge of the world in which she lives and the deliberative capacity to achieve her final desires in it. A fully rational and coherent agent has to desire these things themselves in order to function properly as an ideal agent. Smith insists that the realization of these two dominant final desires is a condition to realize all other possible idiosyncratic desires that an agent's rational counterpart might happen to have.

But the problem now is that, even though all ideal agents have those two dominant final desires of the ideal agency, those desires are restricted to help to ensure and not to interfere with the present and future exercise of one's own capacities. So there is no guarantee that they all will converge on the same set of desires over time and space. To solve that, Smith holds the *symmetry argument* according to which an agent's relationship to her future self is not different from her relationship with other people. The core idea is that an ideal agent would desire to help to ensure and not to interfere with the current and future exercise of the two deliberative capacities of not just herself, but of anyone whose exercise of their two deliberative capacities is dependent on what she currently does.

On this view, an agent must not give particular weight to what she now wants or values. She must give equal weight to all the parts of her life, or to what he wants or values at all times. Further, it would be arbitrary for agents to treat themselves and other differently when they all have similar interests in acting rightly. To use Smith's (2013, p. 24) words,

[t]hrough there is a deep difference between an agent's relationship to her own current beliefs and desires, and those she has later, [...] there is no such a deep difference between her relationship to her own later beliefs and desires and those of other people. [...]. Just as an agent's being ideal at a time requires her to be, at that time, such as she needs to be in order to be ideal, not just at that time, but also at the later times, so her being at a time requires her to be, at that time, such as she needs to be in order to be ideal not just herself, at that time and at later times, but also as she needs to be for others to be ideal, whether at that time or at later times.

The second part of the constitutivist approach regards the normative significance of the ideal agency. In general terms, Smith argues we should understand reasons for action in terms of what is constitutive of the agency, where this means to reduce facts about the wrongness of actions to facts about the rational requirements to which actions are subject. It is an attempt to establish moral conclusions by saying that an ideal deliberator must have certain dominant final desires because she is ideal. On this view, having these dominant final desires is required of agent's rational counterpart to meet the highest standards that are internal to the concept of agency. Thus, if some final

desires are constitutive of the agency, then all ideal agents will have them under conditions of full rationality and coherence.

According to Smith's account, it is possible to reduce moral requirements to rational requirements insofar as one has a moral reason to desire a given action just in case she would desire it if she maximally exercised the two deliberative capacities of herself and other people, which are constitutive of the agency. Thereby, helping to ensure and not interfering with the exercise of the two deliberative capacities of ourselves and other people is what is morally required us to do at the most fundamental level. In Smith's (2011, p.360) words, as every agent's ideal "counterpart has these desires, every agent has the same reasons for action", where these reasons for action are reasons to do what a fully rational and coherent agent would desire to do¹⁰.

The third and last part of Smith's constitutivism regards the attempt to combine a given conception of ideal agency with the normative significance of it in order to derive some subsidiary moral principles, which are, at least on his view, constitutive of the agency as well. The key idea is that a maximally coherent and unified exercise of the deliberative capacities to know the world in which we live and to achieve our final desires in it suggests that there will be many other moral requirements beyond (a) and (b). Smith (2013, p. 27) points out there are subsidiary moral principles in a sense that not to fulfill them frequently involves disappointing the reasonable expectations of others, thus not helping to ensure and interfering with the exercise of their *knowledge acquisition* and *instrumental reasoning* capacities. So an ideal agent would desire not just to help to ensure and not to interfere with the exercise of the two deliberative capacities, but also not lying, not manipulating, not cheating, not being disloyal, not betraying and so on, given that these actions would undermine the reasonable expectations of others.

Let me try to make this point with Smith's (2013, p. 25) example. Suppose a situation in which I promise to find you around at 5:00 p.m. in front of the cinema to watch a movie and you appear at the agreed time, but I fail to show up. Suppose also that this happens not because my car had a mechanical problem or because I was helping a victim of a traffic accident, but because I did not feel like going when the time came. I knowingly failed to keep that promise without taking the necessary means to warn you and without having some compelling reason to break the promise. Given Smith's argument, what is the problem with what I did?

¹⁰ Obviously, people can fail in the exercise of these two dominant final desires and therefore act wrongly. However, they are responsible for any misperceptions insofar as they had the two deliberative capacities, but failed to rightly exercise them.

The problem is that I knowingly interfere with your exercise of your deliberative capacity to have knowledge of the world in which you live insofar as I led you to believe I would watch the movie with you. Further, I also knowingly interfere with your exercise of your deliberative capacity to achieve your desires insofar as I led you to believe you would act in accordance with your desire to spend the evening with me. If you had not falsely believed that I would be there, you could have acted on some other desire, a desire that you could have had a chance of achieving. Thus, the requirement to keep promises is an example of a subsidiary moral principle in a sense that, when I broke my promise, I undermined your reasonable expectations. Just as the subsidiary moral principles of not lying, not manipulating, not cheating, not being disloyal and not betraying are founded in the fact that the creation of reasonable expectations is intrinsic to human nature, so the subsidiary moral principle of keeping promises is as well.

Smith further points out there is also another sort of subsidiary moral principles. For instance, a father who fails to provide his child with the necessary and available information to form her beliefs about the world in which she lives and who fails to provide her with resilience and self-confidence to persist and achieve her final desires is making a mistake. The development of the child's capacities to know the world in which she lives and get her final desires in it is dependent on what her father does to ensure a maximally fruitful exercise of them in the future. In that way, Smith (2013, p. 28) goes on, the father "was subject to a subsidiary moral requirements to teach her that books are full and useful information, that many people outside the family are to be trusted, that there are many ways for girls to get ahead without being attached to some powerful man, and so on".

In conclusion, the point worth noting here is that Smith develops a constitutivist approach in order to offer a reason to believe that a sufficient amount of convergence in fully rational and coherent agents' desires will emerge over time and space. He believes that the dominant final desires to help to ensure and not to interfere with the exercise of *knowledge acquisition* and *instrumental reasoning* capacities are constitutive of the ideal agency, and therefore they are moral requirements, which is to say, they are actions we are obligated to do. All ideal agents thus would converge on the same set of desires because having these desires is part of having an ideal psychology.

3. Two Objections to the Constitutivist Approach

As I tried to show in the last section, Smith's constitutivist approach rests on the idea of providing a rational foundation for morality by reducing moral requirements to rational requirements. He states that rational

requirements should be understood in terms of dominant final desires of the ideal agency, which tell us what is required of us to do under certain idealized conditions. However, I suspect that Smith has failed to perform such task, and hence his constitutivist account cannot provide a reason to believe why rational and coherent agents will converge over time and space on a common set of desires.

With that in mind, now I intend to discuss more thoroughly two main reasons presented in a general way by Michael Bukoski (2016) for that conclusion. The first of these reasons is that Smith's justification for the *symmetry argument* either rests on a question-begging argument or there is none. In either case, the *symmetry argument* lacks rational justification, and therefore the substantive convergence is not possible insofar as other's interests would not matter to any particular agent. And the second reason is that, even if we accept that Smith can successfully reply the first objection, the problem now is that his view on rational requirements can differ significantly from moral requirements, undermining then the reduction of morality to rationality.

3.1 First Objection

About the first objection, it strikes me that Smith needs to provide a reason to justify why it is arbitrary for agents to treat themselves and others differently, that is, why others' interests and claims cannot be ignored by each agent. Such a reason is needed because without it, no substantive convergence will emerge over time and space. Ideal agency would be compatible with indifference to others' desires and claims, and so we will not get moral objectivity. However, it is worth noting that Smith cannot provide any reason for that. He should provide a reason that does not rely on moral premises, on pain of threatening to beg the question in favor of the rationality of moral requirements by incorporating implicitly moral premises that are not justified by what is constitutive of the ideal agency. So it is imperative for Smith to give a rational foundation for the *symmetry argument*.

A first path Smith appears to go is that he does not provide any rational justification to account for why it would be arbitrary to distinguish ourselves from others or why our relationship to other people is not relevantly different from our relationship to our future self. He believes that if I have a given interest by x and you have a qualitatively identical interest by x , then I ought to treat your and my interests by x as equally important. Nonetheless, I understand that merely postulating the *symmetry argument* is not a justification for holding that it is, indeed, a principle of rationality. At first sight, no one will be convinced of it if there is no rational argument to justify why I ought to treat your and my interests by x as equally important.

In one sense I suspect Smith does not give such a rational justification because by doing so he would be in disagreement with his theoretical framework as a whole. As Smith's conception of rationality is based on the instrumental view according to which what is rational to do just depends on one's desires rather than others', it remains unclear why would be arbitrary to give more attention and strength to my own desires rather than the others'. For even though my and others' interests are similarly constituted, my interests are still mine and others' are still theirs. Thus, pending other arguments for believing in the *symmetry argument*, the mere fact that you have the same interest as I have does not mean that it is rationally arbitrary for me to give more attention and strength to my own interest.

Smith, of course, could reply this by saying that, even though what is rational to do just depends on one's desires, it would be more systematic to expect that fully rational and coherent agents treat similar cases in a similar way. It is constitutive of the ideal agency to have the capacity of being symmetric over similar circumstances. For example, someone can have a more general interest to care about the welfare of other people than merely caring about her own welfare, since other people's welfare is similar to her own.

But the problem then is that it might be just as systematic to act for one's own welfare in all the circumstances, or it might be just as systematic not to care about welfare at all and so prefer some other thing, given that systematicity is only a formal criterion. After all, what makes the interests matter is the content of them and not just the symmetry among them. The point here is that Smith's constitutivism seems to be unable to explain what would be more systematic to care about the welfare of other people rather than something else. Hence, in agreement with this first interpretation of Smith's constitutivism approach, I believe he has supported a given characterization of the ideal agency without providing a rational justification for the *symmetry argument*.

Anyway, for the sake of my argument let's imagine that that interpretation above is wrong. So a second path Smith seems to go is to ground the *symmetry argument* on the premise that every agent has equal moral importance, ensuring thus the impartiality between one's desires and others' (BUKOSKI, 2016, p. 132-3). In such a case, it would be arbitrary for agents to treat themselves and others differently because each of them has equal value and, in that sense, deserves equal treatment. All agents count as one regardless of which are their personal desires. Thereby, it would be morally required of each ideal agent's desires to help to ensure and not to interfere with the current and future exercise of the two deliberative capacities of not just herself, but of

anyone whose exercise of their two deliberative capacities is dependent on what she currently does.

However, such an idea leaves it unclear how moral requirements could be reduced to rational requirements. Recall that Smith cannot provide any sort of reason, since for keeping the plausibility of his argument he needs to give a rational foundation to the *symmetry argument*. But the issue is that he is precisely giving a moral character to it and so threatening to beg the question in favor of the rationality of moral requirements. Then my point here is that Smith does not succeed in establishing rational requirements to help to ensure and not to interfere with the two deliberative capacities of oneself and others because the justification for such an idea rests on a moral premise.

To put the issue another way, in order to explain why agent's relationship to other people is not relevantly different from our relationship to our future self, Smith incorporates on the argument implicitly moral premises that are not justified by what is constitutive of the ideal agency. He incorporates the idea that all agents have equal moral importance and so they deserve equal treatment. Hence, in accordance with this second interpretation of Smith's constitutivism, I believe that, without some other non-question-begging justification for the *symmetry argument*, he cannot explain why it would be arbitrary each agent not to give equal weight to her desires as to those of others.

In short, the key point concerning this first objection is that Smith's *symmetry argument* is to make a move to ensure that others' interests and claims matter to any particular agent. Yet I hold it lacks rational justification in both paths that Smith seems to go. On the one hand, it is so because he provides no rational justification for it. On the other, because its justification rests on a question-begging argument. Without such a rational justification, we will not achieve the substantive convergence insofar as the ideal agency would be compatible with indifference to other's desires and claims. Therefore, if what I have said thus far makes sense, then Smith's constitutivist approach has failed to show how ideal agents would converge on the same set of desires under conditions of full rationality and coherence.

3.2 Second Objection

But now let us suppose that Smith can successfully reply the first objection and so provide a rational justification for the *symmetry argument*. In this sense, the second objection is that his account of rational requirements seems to conflict with morality on certain circumstances. According to Bukoski (2016, p. 137), we can imagine some occasions where Smith's

reduction of morality to rationality is threatened to a greater or lesser degree. Thus, I will try to give one example where such a reduction could have some counterintuitive normative implications.

According to Smith's (2013, p. 26) argument, the reduction of morality to rationality is possible the extent to which rational requirements has a striking similarity to what "we ordinarily take to be morally required" to do. He justifies the existence of some subsidiary moral principles by saying that not to fulfill them typically involves disappointing the reasonable expectations of others, thus not helping to ensure and interfering with the exercise of the two deliberative capacities of the ideal agency. Smith (2013, p. 27) believes that the existence of these subsidiary desires is founded in the fact that the "creation of reasonable expectations is ubiquitous in human nature". From that outlook, an ideal agent would desire, for example, not lying, not manipulating, not cheating, not being disloyal, not betraying and not breaking promises, for these actions would undermine the reasonable expectations of other people. But I suspect that this claim is problematic on some occasions.

My argument is as follows. If the rational requirement that supports the existence of these subsidiary moral principles is that the realization of certain actions would undermine one's reasonable expectations and so not help to ensure and interfere with the exercise of his two deliberative capacities, then this rational requirement could make "moral requirements where there are none" (BUKOSKI, 2016, p. 137). It would be possible to derive some very odd moral conclusions from such rational requirement in virtue of the fact that our habits and daily actions at times create reasonable expectations in other people. But it does not follow from that we are morally wrong to disappoint these expectations by acting in another way.

For example, imagine a situation in which I am in love with a girl, and she does not know about my feelings for her. I know she goes alone every Wednesday at the cinema, and then I decide to go to the cinema on a given Wednesday to find and tell her about my feelings. Nonetheless, suppose that in the exact Wednesday I go to the cinema to do that she is not there. In such a case, she would undermine my reasonable expectations and so not help to ensure and interfere with the exercise of my two deliberative capacities. From Smith's reduction, we should say she is doing something morally wrong. But I believe it is plainly implausible to conclude from that she is morally required to go to the cinema if she wants to do something else instead.

As much as I can see, the example above brings one occasion where Smith's justification for the existence of some subsidiary moral principles has some counterintuitive implications. It strikes me that reasonable expectations can be undermined without moral wrongness. But it is worth noting here I am

not claiming that breaking a promise or being disloyal are not morally reprehensible things to do. My point is just to claim that the rational justification that Smith provides for a certain group of subsidiary moral principles, if applied to another range of circumstances, it could lead to a group of quite strange moral conclusions. If this makes sense, then rational requirements may differ significantly from moral requirements, and thus Smith's reduction fails to account for why some dominant final desires are constitutive of the ideal agency – i.e., proposition (2).

Conclusion

The aim of this paper has been to answer the following question: Does Smith's theory account for how agents would ideally converge on the same set of desires over time and space? Smith has defended a sort of constitutivism according to which there are some desires which are constitutive of the ideal agency. On his view, all agents would have the same set of desires because having these desires is part of having an ideal psychology. Yet I have tried to suggest we should respond the question above negatively, for Smith's constitutivist is not capable of providing us a reason to believe in the substantive convergence among agents' desires under certain idealized conditions. In support of that response, I have claimed his approach is victim of two objections in particular. First, by trying to ensure that others' interests and claims matter to any particular agent, Smith's *symmetry argument* lacks rational justification. And second, by trying to derive some subsidiary moral principles from the idea that not to fulfill some moral principles involves disappointing the reasonable expectations of others, his theory makes moral requirements where there are none. Hence, I conclude by saying that Smith's constitutivism has failed to explain how moral judgments express the subject's belief about an objective matter of fact – i.e., proposition (1).

References

- BUKOSKI, M. "A Critique of Smith's Constitutivism". In: *Ethics*, v.127, n.1, 2016, p.116-146.
- COPP, D. "Belief, Reason, and Motivation: Michael Smith's 'The Moral Problem'". In: *Ethics*, v.108, n.1, 1997, p.33-54.
- DWORKIN, D. *Taking Rights Seriously*. Cambridge: Harvard University Press, 1977.
- ENOCH, D. "Rationality, Coherence, Convergence: A Critical Comment on Michael Smith's 'Ethics and the A Priori'". In: *Philosophical Books*, v.48, n.2, 2007, p.99-108.

- FERRERO, L. "Constitutivism and the Inescapability of Agency". In: SHAFER-LANDAU, R. *Oxford Studies in Metaethics 4*. New York: Oxford University Press, 2009.
- FOOT, P. "Morality as a System of Hypothetical Imperatives". In: *The Philosophical Review*, v.81, n.3, 1972, p.305-316.
- JOYCE, R. *The Myth of Morality*. Cambridge: Cambridge University Press, 2001.
- KIMLICKA, W. *Liberalism, Community and Culture*. New York: Clarendon Press, 1989.
- KORSGAARD, C. *Self-Constitution: Agency, Identity, and Integrity*. New York: Oxford University Press, 2008.
- MILLGRAM, E. *Varieties of Practical Reason*. Cambridge: MIT University Press, 2001.
- RAWLS, J. *A Theory of Justice*. Cambridge: Harvard University Press, 1971.
- SMITH, M. *The Moral Problem*. Oxford: Wiley-Blackwell, 1994.
- _____. "Deontological Moral Obligations and Non-Welfarist Agent-Relative Values". In: *Ratio*, v. XXIV, 2011, p.351-63.
- _____. "A Constitutivist Theory of Reasons: Its Promise and Parts". *LEAP: Law, Ethics, and Philosophy*, (1), 2013, p.9-30.
- _____. "The Magic of Constitutivism". In: *American Philosophical Quarterly*. (52), 2015, p.187-200.
- THOMPSON, J. *Normativity*. Chicago: Open Court, 2008.
- VELLEMAN, D. *The Possibility of Practical Reason*. Ann Arbor: Maize Books, 2009.
- WILLIAMS, B. *Moral Luck*. Cambridge: Cambridge University Press, 1981.

E-mail: mdalsotto@hotmail.com

Recebido: 22/06/2017

Aprovado: 23/08/2018