

# Justice as Coherency

## *Reflective Equilibrium and the Justifying Process*

Anna Rus

5653886

Department of Philosophy

Faculty of Humanities

September 28, 2018

Supervisor: Sem de Maagt

Second Assessor: Niels van Miltenburg

# Index

Summary .....	3
1. Introduction .....	4
2. A Theory of Justice .....	5
<b>2.1 A Theory of Justice: Original Position</b> .....	5
<b>2.2 A Theory of Justice: Reflective Equilibrium</b> .....	6
<b>2.3 Wide Reflective Equilibrium</b> .....	8
3. Coherency as Justification .....	9
<b>3.1 Is Coherency Enough?</b> .....	9
<b>3.2 Provisional State of Affairs</b> .....	12
4. Possible Solutions? .....	13
<b>4.1 Adjusting the theory: “politicizing” <i>Justice as Fairness</i></b> .....	13
<b>4.2 Misrepresenting Analogy</b> .....	17
<b>4.3 Expanding the Method</b> .....	18
5. Justification through Coherency: A fundamental Problem .....	20
6. Conclusion .....	21
Bibliography .....	23

## Summary

This paper analyses the method of reflective equilibrium as an account for justification. John Rawls presented this coherentist view in his '*A Theory of Justice*' as an alternative for foundationalist approaches to justification, in order to avoid the epistemological and metaphysical problems of these approaches. Without making controversial claims about the metaphysical status of moral principles and judgments, Rawls claimed that reflective equilibrium could offer justification by producing coherency between beliefs about the issue at hand.

This paper will examine the question whether or not the method of reflective equilibrium can account for justification as claimed, and will answer this question with 'no'. This answer will be found by analyzing the method of reflective equilibrium, showing that the method needs justified input to generate justified output. In other words, for the produced coherency to be justified, the input of the method needs to be justified *before* brought into reflective equilibrium. As a consequence, the principles following from the method of reflective equilibrium might very well be justified, however this justification is not provided by the method itself.

In order to find out if this problem can be overcome, the paper will focus on some of the arguments defenders of the method of reflective equilibrium have presented as well. It will be shown that they offer a better understanding of the method, but that their solution lies in adjusting the input of the method as well: leaving the problem of the lack of justifying force in the method unsolved.

It will be concluded that reflective equilibrium as a justifying method seems to be an attractive device because of its epistemological and metaphysical neutrality, but that it merely produces coherency which cannot provide justification. For principles to be justified, the input of the method needs to be justified prior to the method of reflective equilibrium: the method itself does not have the resources to justify.

# 1. Introduction

In 1972, John Rawls published his work '*A Theory of Justice*'. The theory he presents us there, with the aim to organize a just, liberal society, is called *Justice as Fairness* and will be his most discussed work. Even today, the theory still has a great influence on political philosophy and questions about just societies. As we will see in this paper, *Justice as Fairness* begins with a thought experiment called 'the original position', in which parties will come to agreement about just principles for the society they live in. These principles will be justified by the method Rawls called *reflective equilibrium*.

With offering the method of reflective equilibrium as an account of justification, Rawls avoids the epistemological and metaphysical problems of the so-called foundationalist approaches. Because his offered method avoids these problems, it is considered a very attractive view. However, even though Rawls avoids the typical problems for the foundationalist views, his method of justification is still subject to a lot of criticism. It is often asked how exercising the method of reflective equilibrium will lead to justified principles: is this even a real possibility?

In this paper, I will argue that the question as to whether or not reflective equilibrium can provide for justification should be answered with 'no'. This has the consequence for Rawls's theory that even though principles can be generated and there might exist coherency between them, there is no guarantee that the principles are actually *justified*, and, therefore, *just*. There are ways that seem to avoid this problem: you can add several conditions to the "input" of reflective equilibrium, so that the "output" will in fact be justified. However, this does not fix the inability of the method of reflective equilibrium to justify.

The structure of the paper will be the following. First, I will introduce the two main concepts as Rawls presented them in *A theory of Justice*, being the original position and reflective equilibrium. I will explain the function of both concepts and the eventual purpose they serve. For a clear understanding of reflective equilibrium, I will focus on an explanation about reflective equilibrium as given by Norman Daniels.

After I have done that, I will explain why I think reflective equilibrium as introduced by Rawls has a fundamental problem in section 3. I will introduce an article written by Kelly and McGrath and use this to show that coherency between principles and judgments is not sufficient for justification of principles.

In section 4, I will present some possible defenses to the problem stated in this paper. First, I will focus on adjustments and changes Rawls has made throughout the years, with an overview of his 'political project' as presented in '*Justice as Fairness: Political not Metaphysical*', and '*Political Liberalism*'. The focus will mainly be on the differences between the views of reflective equilibrium between *A Theory of Justice* and *Political Liberalism*. I will show that even though Rawls made critical changes that make his theory definitely more realistic, they however do not solve the problem as posed before.

In what follows, I will discuss two other defenders of reflective equilibrium and examine whether or not they offer a solution for the 'coherency-problem' as I stated it in this paper. In the last section, I will conclude that even though some of the concepts are clarified, the problem of reflective equilibrium still exists. Reflective equilibrium does not have the recourse to account for justification of principles.

## 2. A Theory of Justice

As mentioned before in the introduction of this paper, Rawls wrote 'A Theory of Justice' in which he develops his theory *Justice as Fairness*. In this book, he talks about the requirements to organize a liberal, just society. Two concepts he introduces in order to achieve this, are the original position and the method of reflective equilibrium.

### 2.1 A Theory of Justice: Original Position

Rawls's goal with the original position is to give a description of a situation in which a fair agreement for all parties concerned is reached. Because all agreements reached in this position are fair, the original position will lead us to *Justice as Fairness*.<sup>1</sup> The principles following from the agreement reached in this position are justified by deliberation between the parties concerned. In order to do so, parties will have to establish a set of principles that is rational to adopt in the original situation.<sup>2</sup> This aspect, the one of establishing a set of principles that would be rational to adopt, connects Rawls's theory of justice with a theory of rational choice: we have to decide which principles are to be seen as just *from a rational point of view*.

The conception of the original position should be "the most philosophically favored interpretation."<sup>3</sup> This interpretation is found in reflective equilibrium, to which I will come later. For now, it is important to see that *Justice as Fairness* consists, in Rawls's words, of two parts: On the one hand we need an interpretation of the initial situation, being the original position, and on the other hand we need a set of principles which are rational to adopt and would be agreed to by the people in the original position.<sup>4</sup>

To rule out any unfairness, the principles agreed to in the original position are chosen behind *the veil of ignorance*. The characteristic of this state is the depriving of the parties of the ability to identify principles that would favor their own particular class or group. The idea is that one is disconnected from his or her place in society, his or her social status or class and no one has any knowledge about the distribution of natural recourses and capacities such as intelligence and strength.<sup>5</sup> According to Rawls, this veil of ignorance is necessary to correct the world from arbitrariness: the restrictions on the information is of fundamental importance, because without this the outcome "is biased by arbitrary contingencies."<sup>6</sup>

Another reason for Rawls to rule out knowledge is that of unanimity. Considering that the agreement in the original position is reached through consensus, this plays an important role. Rawls argues that without knowledge of the above described subjects, it is not misplaced to require unanimity because it can be reached under the given circumstances: this possible satisfaction of unanimity is of "great importance."<sup>7</sup> The knowledge that the parties do have, consists of the general facts about human society and the fact that their society is subject to the circumstances of justice.<sup>8</sup> It is important to note the condition of the *veil of ignorance*

---

<sup>1</sup> John Rawls, *A Theory of Justice*, Revised Ed., (Cambridge: Harvard University Press, 1999), 15.

<sup>2</sup> Rawls, *A Theory of Justice*, 16.

<sup>3</sup> Rawls, *A Theory of Justice*, 16.

<sup>4</sup> Rawls, *A Theory of Justice*, 14.

<sup>5</sup> Rawls, *A Theory of Justice*, 11.

<sup>6</sup> Rawls, *A Theory of Justice*, 122.

<sup>7</sup> Rawls, *A Theory of Justice*, 122.

<sup>8</sup> Rawls, *A Theory of Justice*, 119.

does not go against the part of rational choice within Rawls's theory of justice, and so the assumption that all parties are rational.<sup>9</sup> The parties know that they should try to protect their liberties and increase their means that will favor their aims. The parties will choose a plan that satisfy more of their desires, rather than less.<sup>10</sup> This claim is supported by general knowledge about moral psychology and the assumption that rational people want more primary goods, because more primary goods generally assure greater success than less primary goods.<sup>11</sup> Rawls claims that parties have enough knowledge to rank alternatives and can make a rational decision in the ordinary sense.<sup>12</sup>

Last, but not least, are the requirements of the principles that are necessary for assigning the basic rights and duties in society. In order to reach agreement about *just* principles, there are certain requirements that the principles need to fulfill. Rawls claims that those conditions are natural, however necessary, because they will exclude egoism which in turn gives the conditions moral force.<sup>13</sup> The conditions, Rawls writes, are justified by the reasonableness of the theory.<sup>14</sup> Rawls writes that: "A conception of right is a set of principles, general in form and universal in application, that is to be publicly recognized as a final court of appeal for ordering the conflicting claims of moral persons."<sup>15</sup> The parties in the original position are rational, the conceptions of justice will have this content of generality, universality, public recognition and a 'last resort' character.<sup>16</sup>

## 2.2 A Theory of Justice: Reflective Equilibrium

As described before, the justification of the principles agreed to as well as the "most favorable interpretation" of the original position, is given by the method of reflective equilibrium. In other words: reflective equilibrium is used as an account for moral justification.

Rawls writes: "I assume that eventually we shall find a description of the initial situation (original position) that both expresses reasonable conditions and yields principles which match our considered judgments duly pruned and adjusted. This state of affairs I refer to as reflective equilibrium."<sup>17</sup>

Rawls offers the method of reflective equilibrium as an alternative to the so-called foundationalist methods of justification. Roughly, the idea of foundationalist methods of justification is that some moral beliefs are justified because they are dependent on an ultimate moral foundation.<sup>18</sup> This ultimate moral foundation leads to several epistemological and metaphysical problems, because it is unclear how exactly we could reach knowledge about this ultimate foundation and what exactly the status of this foundation is. This makes the position unattractive.

---

<sup>9</sup> Rawls, *A Theory of Justice*, 123.

<sup>10</sup> Rawls, *A Theory of Justice*, 124.

<sup>11</sup> Rawls, *A Theory of Justice*, 79.

<sup>12</sup> Rawls, *A Theory of Justice*, 123.

<sup>13</sup> Rawls, *A Theory of Justice*, 113.

<sup>14</sup> Rawls, *A Theory of Justice*, 113.

<sup>15</sup> Rawls, *A Theory of Justice*, 117.

<sup>16</sup> Rawls, *A Theory of Justice*, 129.

<sup>17</sup> Rawls, *A Theory of Justice*, 18.

<sup>18</sup> Sem de Maagt, "Reflective Equilibrium and Moral Objectivity", *Inquiry: An Interdisciplinary Journal of Philosophy* 60, no. 5, (July 4<sup>th</sup>, 2017): 447.

In the method of reflective equilibrium, on the other hand, justification depends on coherency between a set of beliefs. Therefore, reflective equilibrium is a coherence account of justification.<sup>19</sup> These beliefs are revisable, and the method makes no claim about directly justified beliefs. This means that there is no special epistemological story that needs to be told about these beliefs, which makes the method of reflective equilibrium neutral with regards to metaphysics and epistemology.<sup>20</sup>

In short, this means that the reached coherency between a set of beliefs is the justifying force of the method.<sup>21</sup> Thus, reflective equilibrium will lead us to a coherent set of principles and judgments that will be agreed upon in the original position and, at the same time, the structure of this position is itself justified by the same method of reflective equilibrium.

Something important to note here, is that Rawls talks about *considered* judgments, rather than judgments in general. The idea of considered judgments is that they are rendered under favorable conditions when we are in *the right state of mind*, or, as Scanlon describes it: “judgments that seem clearly to be correct under conditions conducive to making good judgments of the relevant kind.”<sup>22</sup> In other words, considered judgments are the judgments in which our moral capacities are most likely to be displayed without distortion: only then the judgments are allowed into reflective equilibrium.

The method of reflective equilibrium is of fundamental importance within *A Theory of Justice*. Rawls admits there are many interpretations of the initial situation possible and even considered judgments may be subject to irregularities and distortions.<sup>23</sup> However, these problems of distortion and varieties are solved through the method of reflective equilibrium: “...the question of justification is settled, as far as it can be, by showing that there is one interpretation of the initial situation which best expresses the conditions that are widely thought reasonable to impose on the choice of principles yet which, at the same time, leads to a conception that characterizes our considered judgments in reflective equilibrium.”<sup>24</sup> Thus, by exercising the method of reflective equilibrium we can rule out any distortion that might occur in our judgments. In reflective equilibrium, we are allowed to revise our moral judgment to eventually end up with justified judgments.

For the justification of the principles of justice that are claimed to follow from the considered judgments, reflective equilibrium knows broadly speaking three stages: the identification of considered judgments, the formulation of the principles that would account for these judgments and the response to the possible divergence between these principles and one’s considered judgments.<sup>25</sup> Therefore, the fundamental justifying strategy of justice in *Justice as Fairness* is to look for an internally coherent set of principles and considered

---

<sup>19</sup> Norman Daniels, “Reflective Equilibrium” in *The Stanford Encyclopedia of Philosophy* (fall 2018 edition), ed. Edward N. Zalta, <https://plato.stanford.edu/archives/fall2018/entries/reflective-equilibrium/> (accessed September 11th, 2018), 1.

<sup>20</sup> De Maagt, “Reflective Equilibrium and Moral Objectivity”, 445.

<sup>21</sup> Daniels, “Reflective Equilibrium”, 1.

<sup>22</sup> Thomas M. Scanlon, “Rawls on Justification”, in *The Cambridge Companion to Rawls*, ed. Samuel Freeman, (Cambridge: Cambridge University Press, 2002), 140.

<sup>23</sup> Rawls, *A Theory of Justice*, 105.

<sup>24</sup> Rawls, *A Theory of Justice*, 105.

<sup>25</sup> Scanlon, “Rawls on justification”, 140-141.

judgments: exactly this is done in reflective equilibrium.<sup>26</sup> According to Rawls, "...the one [account] which matches his judgments in reflective equilibrium" is the best account of a person's sense of justice.<sup>27</sup> The method of reflective equilibrium should have us arrived at objective principles of justice.<sup>28</sup> *Justice as Fairness*, then, is "a theory of our moral sentiments as manifested by our considered judgments in reflective equilibrium."<sup>29</sup> Our conception of justice we agreed upon is justified, if this conception is the most acceptable one we find in exercising the method of reflective equilibrium.<sup>30</sup>

### 2.3 Wide Reflective Equilibrium

Because Rawls keeps his explanation about reflective equilibrium in *A Theory of Justice* rather short, I will elaborate on the method a little further by focusing on an article written by Norman Daniels, called "*Wide reflective equilibrium and theory acceptance in ethics*". Daniels talks about the same method as meant by Rawls in *A Theory of Justice*, but gives a more comprehensive and elaborated explanation about it. After this, I will continue with what I think is the fundamental problem of the method of reflective equilibrium in section 3.

The definition Daniels gives is: "The method of reflective equilibrium is an attempt to produce coherence in an ordered triple of sets of beliefs held by a particular person, namely, (a) a set of considered moral judgments, (b) a set of moral principles and (c) a set of relevant background theories."<sup>31</sup> The eventual goal is that the principles found in (b) are the most acceptable principles out of alternative ones, independent from their coherency with the considered moral judgments in (a), shown by the relevant background theories in (c).<sup>32</sup>

In order to reach this goal successfully, wide reflective equilibrium permits that the moral judgements are revised in an extensive way.<sup>33</sup> The method prevents us from taking the considered judgments at face-value. Instead, we are continually judging which of the considered moral judgments needs revision: it is not a mere systemizing of a determinate set of judgments.<sup>34</sup> Daniels compares the method of wide reflective equilibrium with scientific practice. He claims that in ethics, the same way as in science, we keep reassessing and reevaluating our data against the accepted theories: there is a constant pressure for revision, rather than a simple test of our theories against a predetermined body of data.<sup>35</sup>

To summarize, the idea that Rawls presents in *A theory of Justice* is that of *Justice as Fairness*. In this conception of justice, free and rational persons agree to a set of principles that assign rights and duties in society. This is done in the original position behind the veil of ignorance.

---

<sup>26</sup>Onora O'Neill, "The Method of a Theory of Justice", in *John Rawls, A Theory of Justice*, (Leiden: BRILL, 2013), 30.

<sup>27</sup> Rawls, *A Theory of Justice*, 46.

<sup>28</sup> De Maagt, "Reflective Equilibrium and Moral Objectivity", 449.

<sup>29</sup> Rawls, *A Theory of Justice*, 104.

<sup>30</sup> Norman Daniels, "Reflective Equilibrium and Justice as Political" in *Justice and Justification: Reflective Equilibrium in Theory and Practice*, (Cambridge: Cambridge University Press, 1996), 153.

<sup>31</sup> Norman Daniels, "Wide Reflective Equilibrium and Theory Acceptance in Ethics", in *Justice and Justification: Reflective Equilibrium in Theory and Practice*, (Cambridge: Cambridge University Press, 1996), 22.

<sup>32</sup> Daniels, "Wide Reflective Equilibrium", 22.

<sup>33</sup> Daniels, "Wide Reflective Equilibrium", 27.

<sup>34</sup> Daniels, "Wide Reflective Equilibrium", 28.

<sup>35</sup> Daniels, "Wide Reflective Equilibrium", 33.



Both aspects of this theory, being the original position and the principles of justice agreed to, are justified through reflective equilibrium. Instead of accepting that a considered judgment should be justified since it is *considered*, we always keep on testing them.

### 3. Coherency as Justification

Having explained Rawls's position, I will now turn to what I think is a very hard to escape problem of his theory and the method of justification presented in it. Even though Rawls claims that his theory of justice is eventually justified by the method of reflective equilibrium, I think this premise is mistaken. The main point is this: reflective equilibrium can only guarantee coherency, but that in itself is *not* sufficient for the justification of principles. This will have serious consequences for Rawls's theory: in order to make sure that the reached coherency actually *does* give a sufficient account for justification, we need to adjust the input of the method. That is, we need to refine our considered judgments and add several conditions to them that will make them acceptable for reflective equilibrium. However, that way the justification of those principles is given *before* reflective equilibrium is even reached. The conditions as stated by Rawls do not suffice for a justification of the principles and even if they did, they will eventually bring us no more than an adjustment of the starting points: leaving the problem of the method of justification unsolved.

#### 3.1 Is Coherency Enough?

The justification of the abovementioned principles of justice that are agreed to is given in reflective equilibrium. As explained, this justification is a method of going back and forth: sometimes the conditions of the original position are altered, and other times our considered judgments are withdrawn in conforming them to principles. Proceeding this way, we will come to an interpretation of the original position that is most favored in which the persons will reach agreement about just principles.<sup>36</sup> As emphasized by Daniels, reflective equilibrium is an attempt to produce coherence.

We have seen that exercising the method of reflective equilibrium does not make moral justification depend on a moral foundation but rather on a coherency between all moral and non-moral beliefs that a person has.<sup>37</sup> Exactly this aspect, namely the production of coherence, is where I think the fundamental problem can be found. The way reflective equilibrium is described by Rawls and, more extensively, by Daniels, it seems to be more like a comparing method that will eventually lead to coherency, rather than a justifying method. To make this coherent set of principles justified, we need more than just reflective equilibrium.

Let me explain this by giving a rather simple example. Assume that we, or in case of *A Theory of Justice* the parties in the original position, start off with an '*unjust*' description of the conditions of the original position. Even though these conditions are unjust, the persons have several considered judgments that are brought into this position. With the method of reflective equilibrium, coherency will be produced between these considered judgments and the conditions and principles of the original position. The outcome will be, even though it is perfectly *coherent*, unjust.

---

<sup>36</sup> Rawls, *A Theory of Justice*, 18.

<sup>37</sup> De Maagt, "Reflective Equilibrium and Moral Objectivity", 447.

The same argument will be true when we reason the other way around. The persons in the original position start off with perfectly just conditions. They bring in several considered judgements, but this time these judgments are unjust. They will, just as in the first example, exercise the method of reflective equilibrium and try to reach coherency. However, instead of withdrawing their unjust judgments, they will adjust the conditions of the original position. Again, the outcome of the method is perfectly coherent but still *unjust*.

This problem is also described in the article "*Reflective Equilibrium and Moral Objectivity*", written by De Maagt. He writes: "It seems that the method of reflective equilibrium cannot provide the resources to exclude these kinds of moral prejudices being brought into coherence with reflective equilibrium. Thus, as long as the input into the method of reflective equilibrium is not initially credible, the equilibrium arrived at after due reflection may be no more than a reshuffling of moral prejudices."<sup>38</sup> This is exactly what I tried to show in the examples stated above: whenever the input of the method is unjust, the outcome will be, however coherent, unjust as well.

To demonstrate that coherency is insufficient to guarantee justification, I will show that the method of reflective equilibrium does not have the resources to exclude unjust judgments and explain why this is a problem for Rawls's theory. I will do so by starting with an explanation of a critique written by Kelly and McGrath in their article "*Is Reflective Equilibrium Enough?*"

Simply put, this section will explain that the problem of the method of reflective equilibrium is that *unreasonable* or *unjust* outcomes of the method are a real possibility, even when the method is executed in the right way. A way to overcome this problem is to adjust the so-called 'starting points' of the method. However, by doing so the method loses its justifying force.

As explained before, in a coherentist view an adequate level of coherency is sufficient for justification.<sup>39</sup> In fact, this is what Rawls says in *A Theory of Justice*: "...justification rests upon the entire conception and how it fits in with and organizes our considered judgments in reflective equilibrium. As we have noted before, justification is a matter of the mutual support of many considerations, of everything fitting together in one coherent view."<sup>40</sup> This means, in other words, that whenever one arrives at a coherent view after due reflection, this view (consisting of moral principles) is justified. Or, in Rawls's words: "a moral judgment is justified for an individual if she holds in in a state of wide reflective equilibrium."<sup>41</sup>

What I will show is that this coherency is *only* enough when there are conditions for the considered judgments as 'starting points' of the method. This in turn will have the consequence that the method does not have any justifying force, but instead justification is given by the conditions for the starting points.

In the article "*Is Reflective Equilibrium Enough?*", Kelly and McGrath give the critique that the method of reflective equilibrium is too weak in the sense that one might arrive at 'hopelessly inadequate' views, even though this person has applied and executed the method

---

<sup>38</sup> De Maagt, "Reflective Equilibrium and Moral Objectivity", 454.

<sup>39</sup> Kelly and McGrath, "Is Reflective Equilibrium Enough?": 332.

<sup>40</sup> Rawls, *A Theory of Justice*, 507.

<sup>41</sup> Kelly and McGrath, "Is Reflective Equilibrium Enough?": 335.

perfectly.<sup>42</sup> The mere objection that can be made that there is a possible situation of following the method of reflective equilibrium and ending with ‘radical error’, is not the main problem. In fact, as Kelly and McGrath say, the possibility of making an error is not a good objection for a theory at all. As in science, we will sometimes make an error simply because the empirical evidence that we use can mislead us and is therefore unrepresentative.<sup>43</sup> However, we do have a serious problem when the theory can, after impeccable execution, lead us to *unreasonable* views to hold: a problem that seems real for the method of reflective equilibrium.<sup>44</sup>

A defense to this might be that this unreasonable ‘output’ is caused by unreasonable ‘input’. It is simply too much to ask that a method can deliver *reasonable* output while presented with *unreasonable* input.<sup>45</sup> Therefore, unreasonable output has to do with the input of the method but not with the method itself. This is also much like the tactic used by defenders of the method of reflective equilibrium and what these defenders say in a disguised manner: there are several conditions in order for starting points to be acceptable. These conditions are necessary to be able to exercise the method of reflective equilibrium.<sup>46</sup>

In their article, Kelly and McGrath are concerned with the question: “why think that the best way for each person to achieve coherence among his or her own judgments will lead to a convergence?”<sup>47</sup>

They examine this with the introduction of two questions. First, the intrapersonal question that asks: “For any particular person, is there some unique reflective equilibrium that she would arrive at if she employed the method impeccably?” And second, the interpersonal question that asks: “Would different individuals, each of whom employed the method impeccably, converge on a unique reflective equilibrium?” They make the following point: even when we assume that the answer to the first question is yes, this is not sufficient for a positive answer to the second question.<sup>48</sup>

The argument they present is that because there are many admissible starting points that will *not* lead to convergence, there is no reason to believe that the coherency between beliefs of one individual will lead to convergence. This convergence is almost impossible because even though there are conditions for the admissible starting points (being the considered judgments), Rawls is extremely liberal in what can actually count as one.<sup>49</sup> About this liberty concerning acceptable starting points, Kelly and McGrath write: “Similarly, given the characterization of considered judgments, it is quite clear that different individuals might begin from radically different sets of judgments, all of which qualify as considered judgments.”<sup>50</sup> Therefore, the possibility of conflicting sets of judgments is high which makes the plausibility for convergence less likely.

Even though Kelly and McGrath focus on the unlikely chances of convergence and the consequences of moral truth, whereas I focus on the justification of the outcome of the

---

<sup>42</sup> Thomas Kelly and Sarah McGrath, “Is Reflective Equilibrium Enough?” in *Philosophical Perspectives* 24 (1): 226.

<sup>43</sup> Kelly and McGrath, “Is Reflective Equilibrium Enough?": 326.

<sup>44</sup> Kelly and McGrath, “Is Reflective Equilibrium Enough?": 326-327.

<sup>45</sup> Kelly and McGrath, “Is Reflective Equilibrium Enough?": 327.

<sup>46</sup> Kelly and McGrath, “Is Reflective Equilibrium Enough?": 328.

<sup>47</sup> Kelly and McGrath, “Is Reflective Equilibrium Enough?": 338.

<sup>48</sup> Kelly and McGrath, “Is Reflective Equilibrium Enough?": 337-338.

<sup>49</sup> Kelly and McGrath, “Is Reflective Equilibrium Enough?": 338.

<sup>50</sup> Kelly and McGrath, “Is Reflective Equilibrium Enough?": 339-340.

method of reflective equilibrium, their argument is still interesting.<sup>51</sup> If we want the reached coherency to be sufficient for justification, whether or not this coherency will lead to convergence, we need to change something about the acceptance of widespread divergence in considered judgments.

The point of justified coherency is something Thomas Scanlon says more about in his article '*Rawls on Justification*'. According to Scanlon, the starting points of reflective equilibrium need to have some initial credibility.<sup>52</sup> However, if this is right, the justifying force lies in the credibility of the considered judgments itself and *not* in the method of reflective equilibrium. Another problem is that in the understanding of Rawls, something might very well be a *considered judgment*, while completely lacking this initial credibility.<sup>53</sup>

In order to solve this problem, Kelly and McGrath offer a characterization of the starting points in a normative way.<sup>54</sup> This will mean that starting points are only accepted if they, for example, are propositions that have a "sufficiently high credibility", or when one is justified in believing *before* used as input for reflective equilibrium.<sup>55</sup> They argue that this is necessary for a defensible account of the method of reflective equilibrium.<sup>56</sup> I think they are right in doing so, since this is the *only* way one can guarantee some *justified* coherency rather than coherency between completely unreasonable or unjust principles. However, with this move the justificatory force of the method of reflective equilibrium seems to have disappeared.

### 3.2 Provisional State of Affairs

What I have tried to demonstrate so far is the following. With *Justice as Fairness*, Rawls offers us a coherentist theory which makes that the coherency between moral judgments coincides with justification. This justification is offered by the method of reflective equilibrium, or at least, so claims Rawls. I explained that the aim of the method of reflective equilibrium is to produce some type of coherency, which in itself offers justification.

However, this method knows a deep-rooted problem. With the help of Kelly and McGrath's critique, I have showed that coherency in itself does *not* justify. If we want the outcome of the method to be justified, we need several conditions for the starting points of reflective equilibrium. Even though Rawls offers us one condition, namely that the starting points of the method of reflective equilibrium should be considered judgments, this does *not* suffice for the type of justification he is aiming for.

In order to produce a justified outcome of the method, the considered judgments need to have at least some initial credibility and a person that holds those judgments need to be justified in believing them *before* using the judgments as input for reflective equilibrium. Yet, with the application of these conditions the justificatory force of the method disappears. As Kelly and McGrath put it: "the most interesting part of the story concerns not the pursuit of equilibrium itself, but rather what makes it the case that certain points are more reasonable than others, and how we manage to recognize or grasp such facts. In that sense at

---

<sup>51</sup> Kelly and McGrath, "Is Reflective Equilibrium Enough?": 340.

<sup>52</sup> Scanlon, "Rawls on justification", 167.

<sup>53</sup> Kelly and McGrath, "Is Reflective Equilibrium Enough?": 346-347.

<sup>54</sup> Kelly and McGrath, "Is Reflective Equilibrium Enough?": 351.

<sup>55</sup> Kelly and McGrath, "Is Reflective Equilibrium Enough?": 351.

<sup>56</sup> Kelly and McGrath, "Is Reflective Equilibrium Enough?": 353.

least, it seems that reflective equilibrium is not enough.”<sup>57</sup> Not only the justificatory force of the method disappears, but in order to be able to produce justified principles, the theory also might also need to assume several beliefs about *justice*. In order to do this, the theory will be biased: it needs several ideas as ‘starting points’ for the conditions of the considered judgments.

## 4. Possible Solutions?

Having explained what my main point of critique is, I will now discuss some possible defenses against it. First, I will turn back to Rawls. After the publication of *A theory of Justice*, he has changed and adjusted his theory, what might change the conceptions of some of the above explained notions. In order to examine whether or not his theory is still subject to my critique, I will explain some of the changes that Rawls made throughout the years. After doing that, I will focus on some other defenders of the method of reflective equilibrium. They offered some possible solutions to the abovementioned critique, claiming that the critique is actually invalid.

I will present an outline of Rawls’s adjustments to his *Justice as Fairness*, starting with an explanation of Rawls’s article “*Justice as Fairness: Political not Metaphysical*”, followed by an explanation of differences between *A Theory of Justice* and Rawls’s later work *Political Liberalism*. In order to do this, I will use the article “*Reflective Equilibrium and Justice as Political*” written by Norman Daniels. For the sake of the argument, I will explain the overall idea of these works rather briefly and mainly focus my attention on the adjustments that affect the method of reflective equilibrium, as well as on the differences between the presentation of reflective equilibrium in the two works. After focusing on Rawls’s work, I will concentrate on a defense given by Norman Daniels, followed by a defense written by Thomas Scanlon.

In the following sections, I will show that Rawls in fact adjusted the starting points of the method of reflective equilibrium, but that he did not solve the problem I have pointed out. I will demonstrate that the responses of Daniels and Scanlon might offer us more understanding about considered judgments, but that they still do not solve the problem of the absence of justifying force of the method of reflective equilibrium.

### 4.1 Adjusting the theory: “politicizing” *Justice as Fairness*

In his article “*Justice as Fairness: Political not Metaphysical*” (hereafter: *Political not Metaphysical*), and his later work “*Political Liberalism*”, Rawls elaborates on and gives us a further explanation about some concepts presented to us in his book *A Theory of Justice*. Rawls gives us a more detailed explanation about some controversial aspects of *A Theory of Justice*, and therefore modifies some of them. I will point out a few of these adjustments and explanations that I think are essential for this paper.

Because this paper discusses on the role (and the connected problem) of the method of reflective equilibrium, I will mostly focus on the differences between reflective equilibrium in *A Theory of Justice* on the one hand, and in *Political Liberalism* on the other. I do this to find out whether or not the problem of the method of reflective equilibrium as posed in this paper is solved by the adjustments, or not.

---

<sup>57</sup> Kelly and McGrath, “Is Reflective Equilibrium Enough?”: 353-354.

In short, taking the changes together will bring us to the following difference between *A Theory of Justice* and *Political Liberalism*. In *A Theory of Justice*, principles are agreed to in the original position. To be justified, these principles should be coherent with our considered judgments in wide reflective equilibrium. This appeal can be seen as an adequacy test for Rawls's theory: we must think about revising the original position and the principles we agreed to, when this coherency is *not* found in wide reflective equilibrium.<sup>58</sup> In *Political Liberalism*, this justifying part of the theory is divided into two parts: "pro tanto justification" and "full justification".<sup>59</sup> Nevertheless, the full justification *still* rests on the same thing as in *A Theory of Justice*: the principles are justified whenever coherent with our considered judgments in wide reflective equilibrium. Thus, the question arises what the essential details of the adjustments made by Rawls are, in order to possibly solve the stated problem of reflective equilibrium.

In *Political not Metaphysical*, Rawls emphasizes two essential points. First, Rawls explicitly points out that his theory is constructed to apply to a *constitutional democracy* and second, *Justice as Fairness* is to be seen as a *political* conception of justice. This means that there are no questions to be raised about the metaphysical status of the theory, and that the question whether or not there is some type of objective moral standard we can know something about is left aside.<sup>60</sup>

The political conception of justice is necessary, Rawls explains, since there is no general moral conception that can provide us with a publicly recognized basis for a conception of justice.<sup>61</sup> To be a workable conception of justice, he says, the conception should allow the diversity that exist in 'existing democratic societies.'<sup>62</sup> This line of reasoning in which the conception of justice is to be seen as *political*, paying respect to existing pluralism within society, is one Rawls elaborates on in his work *Political Liberalism*'. In *Political Liberalism*, Rawls acknowledges that we need to accept *reasonable pluralism* being a basic fact of the political life.<sup>63</sup> Therefore, *Justice as Fairness* as presented in *A Theory of Justice* needed the 'political revision' so it would eventually become an accurate (political) conception of justice.<sup>64</sup>

The main question that arises with this political conception of justice, is how to find a shared basis for the most acceptable conceptions for justice.<sup>65</sup> Rawls's answer to this question is a familiar one: the conception should "be in accordance with our considered convictions, at all levels of generality, on due reflection."<sup>66</sup> Hence, justification is still given by the method of reflective equilibrium.

Again, the idea of *Justice as Fairness* is "that of a society as a fair system of co-operation between free and equal persons", in which this society is a democratic one.<sup>67</sup>

---

<sup>58</sup> Daniels, "Reflective Equilibrium and Justice as Political", 161.

<sup>59</sup> Daniels, "Reflective Equilibrium and Justice as Political", 162.

<sup>60</sup> John Rawls, "Justice as Fairness: Political not Metaphysical", in *Communitarianism and Individualism*, red. Shlomo Avineri and Avner deShalit, (Oxford: Oxford University Press, 1992), 186.

<sup>61</sup> Rawls, "Justice as Fairness", 188.

<sup>62</sup> Rawls, "Justice as Fairness", 188.

<sup>63</sup> Daniels, "Reflective Equilibrium and Justice as Political", 147.

<sup>64</sup> Daniels, "Reflective Equilibrium and Justice as Political", 148.

<sup>65</sup> Rawls, "Justice as Fairness", 192.

<sup>66</sup> Rawls, "Justice as Fairness", 192.

<sup>67</sup> Rawls, "Justice as Fairness", 195.

Moreover, persons participate fully in this system of co-operation and two moral powers are ascribed to them: they have “a capacity for a sense of justice and a capacity for a conception of the good.”<sup>68</sup> This will give the persons the ability to understand, apply and act from the conception of justice, and to revise, form and pursue in a rational way the conception of the good.<sup>69</sup>

In *Political Liberalism*, Rawls discusses two questions, while considering the existing pluralism in society. The first question concerns legitimacy and asks: “How can it be legitimate to coerce all citizens to follow just one law, given that citizens will inevitably hold quite different worldviews?” The second question, concerned with stability, is why a citizen would willingly obey a law if it imposed on her by a collective body, many of whose members have beliefs and values quite dissimilar to her own.<sup>70</sup> What Rawls tries to accomplish with asking and eventually answering these questions, is fixing the context and starting points for *Justice as Fairness*. The idea of *Justice as Fairness* is said to be the shared elements of our political life in a democratic culture.<sup>71</sup>

In this political conception of *Justice as Fairness*, the theory is to be seen as a free-standing module with its own principles and reasons. Being a complete module, it should give reasonable answers to questions about justice.<sup>72</sup> The justification for the answers given in this module appeals to reasons contained in the public view: this is what Rawls calls “pro tanto justification”. The stability question is answered by reaching “overlapping consensus” for the right reasons.<sup>73</sup> Rawls rejects the idea that there is convergence on a (unique) wide reflective equilibrium: overlapping consensus is not enough to reach one *single* wide reflective equilibrium.<sup>74</sup> Finally, *Justice as Fairness* is justified when it is acceptable in the wide reflective *equilibria* that agents can achieve.<sup>75</sup>

The difference between *A Theory of Justice* and *Political Liberalism* is the way agents fit *Justice as Fairness* in their overall systems of moral beliefs.<sup>76</sup> Still, the key to justification remains wide reflective equilibrium: “justice as fairness is justified for an individual if it is the conception of justice that is the most acceptable to her, given all her other beliefs, in wide reflective equilibrium.”<sup>77</sup> This means that the key element of justification still rests on due reflection as is done in reflective equilibrium.<sup>78</sup> Here, the question arises: if the justification rests on the same method, what, then, are the differences made in this method?

The first thing we have to focus on is the use of the veil of ignorance, in which persons are distracted from several types of knowledge that seem irrelevant to the question of justice.<sup>79</sup> This is important, because the knowledge that is left behind is the knowledge we can form the so called *considered judgments* about, being the starting points for reflective

---

<sup>68</sup> Rawls, “Justice as Fairness”, 197.

<sup>69</sup> Rawls, “Justice as Fairness”, 198.

<sup>70</sup> Leif Wenar, “John Rawls”, in *The Stanford Encyclopedia of Philosophy* (spring 2017 edition), ed. Edward N. Zalta, <https://plato.stanford.edu/archives/spr2017/entries/rawls/> (accessed June 20<sup>th</sup>, 2018), 3.1.

<sup>71</sup> Daniels, “Reflective Equilibrium and Justice as Political”, 148.

<sup>72</sup> Daniels, “Reflective Equilibrium and Justice as Political”, 148.

<sup>73</sup> Daniels, “Reflective Equilibrium and Justice as Political”, 149.

<sup>74</sup> Daniels, “Reflective Equilibrium and Justice as Political”, 149-150.

<sup>75</sup> Daniels, “Reflective Equilibrium and Justice as Political”, 149.

<sup>76</sup> Daniels, “Reflective Equilibrium and Justice as Political”, 150.

<sup>77</sup> Daniels, “Reflective Equilibrium and Justice as Political”, 153.

<sup>78</sup> Daniels, “Reflective Equilibrium and Justice as Political”, 149.

<sup>79</sup> See section 2.1.

equilibrium.<sup>80</sup> As Daniels points out, the difference is that in *A Theory of Justice*, the distinction between relevant and irrelevant features seems to be based on an extensive liberal view whereas in *Political Liberalism*, the boundary between relevant and irrelevant is no more than a reference to features that are irrelevant for political argument – there is no moral or political view that has the overtone in this distinction.<sup>81</sup>

The other differences that we can find between *A Theory of Justice* and *Political Liberalism* are more closely related to the method of reflective equilibrium itself. The first difference is the motivation of the parties involved: in *A Theory of Justice*, Rawls claims that this motivation finds his basis in shared philosophical arguments, whereas in *Political Liberalism*, this motivation is specific for each wide reflective equilibrium. Second, Rawls adds a condition to the wide reflective equilibria contained in the overlapping consensus: they should each include beliefs about reasonable pluralism. This means that each person will accept and understand the implications of reasonable pluralism: it is understood and accepted that people might hold different views to a particular subject. Third, Rawls draws a distinction between political and nonpolitical moral beliefs in *Political Liberalism*, a boundary that also needs to be included in reflective equilibrium. This should specify the priority that needs to be given to political beliefs over nonpolitical beliefs.<sup>82</sup>

As stated before, the only way coherency between principles and considered judgments can be justified, is by adjusting the starting points of the method, *before* they are taken into the method itself.<sup>83</sup> What we can see here is that Rawls does exactly this in his political conception of justice: he specifies the circumstances of the parties and applies his theory to a democratic society. Doing this, he adjusts the starting points of the parties: he adds conditions to what is brought into reflective equilibrium and the features parties have knowledge about. Rawls acknowledges the fact that it is unlikely to believe there is only *one* reflective equilibrium and, therefore, there are different reflective equilibria for all persons.

Summarized, this means that the adjustments made by Rawls have a few consequences for the method of reflective equilibrium, although they do not offer an effective solution to the problem this paper writes about.

With the political conception of justice, starting points of the method of reflective equilibrium have been adjusted. The knowledge about features of the agents in reflective equilibrium has changed from a liberal point of view to simply eliminating irrelevant features for a political discussion. On top of this, Rawls added the condition to the reflective equilibria achieved by the parties that each equilibrium should include beliefs about reasonable pluralism. Last, but not least, an account of the distinction between political and nonpolitical beliefs is included. This account priorities to a certain amount the political beliefs and should be included into each reflective equilibrium as well. These changes all affect the starting points of the method of reflective equilibrium, since they concern the features of the things the parties subject of the theory have *considered judgments* about.

However, this does *not* solve the problem of the *justifying* force of the method. The considered judgments of the parties, now consisting of all possible features that *can* be

---

<sup>80</sup> See section 2.2.

<sup>81</sup> Daniels, "Reflective Equilibrium and Justice as Political", 155.

<sup>82</sup> Daniels, "Reflective Equilibrium and Justice as Political", 156-159.

<sup>83</sup> See section 3.2.



relevant to political discussion, *still* need initial credibility and justification to make sure the outcome of reflective equilibrium is justified.

Admittedly, the changes made by Rawls including the acceptance of reasonable pluralism makes his *Justice as Fairness* much more realistic than presented as in *A Theory of Justice*. However, even after the changes, the problem of the method of reflective equilibrium still exist. That is: reflective equilibrium can only guarantee coherency, but it cannot account for justification.

## 4.2 Misrepresenting Analogy

Having examined and concluded that Rawls's theory is, however the adjustments, still subject to the critique stated in this paper, I will now turn to two other defenders of the method of reflective equilibrium: Norman Daniels and Thomas Scanlon.

Both Daniels and Scanlon defend the method of reflective equilibrium against the critique that considered judgments need *initial credibility* and justification *before* the method, that in turn will lose the justificatory force. According to them, this critique is based on misconceptions about the method.

In an article I mentioned earlier in this essay, "*Wide reflective equilibrium and theory acceptance in ethics*", Norman Daniels gives a more elaborated explanation about how to understand wide reflective equilibrium as presented by Rawls. Furthermore, he defends reflective equilibrium against some of the much-argued problems of the method. I will focus here on the defense that he gives against the critique that the starting points used for the method are inappropriate, consisting of the claim that considered judgments are not foundational enough.<sup>84</sup>

The main idea behind this critique is much in line with the problem I discuss in this paper, namely that the starting points of the method need some initial credibility to lead to a credible and justified outcome of reflective equilibrium. Even though we believe in our considered judgments, they might still be incredible.<sup>85</sup>

Daniels believes this problem follows from a misunderstanding about considered judgments, because of the assumption that considered moral judgments and observation reports are analogue.<sup>86</sup> However, he claims, moral judgments like justice and injustice are far more complex than simple observations. Because of that, they probably do not play a role similar to observations, meaning that credence does not have to imply credibility.<sup>87</sup> Thus, Daniels argues that we should not compare moral judgments to observation reports because they are of a different kind. Even though we can add initial credibility to observations, we cannot do the same with moral judgments.

However, distinguishing moral judgments from observation reports does not fix the problem. In fact, the note that we *cannot* add initial credibility to considered judgments only makes the method more troubling. If we cannot reason *if* a moral judgment has initial credibility, it will even be more challenging to reason *why* a considered judgment is justified.

Also, according to Daniels, the "no credibility" problem is 'at best premature'. We are not able to assign initial credibility to a moral judgment, he claims. What this problem shows,

---

<sup>84</sup> Daniels, "Wide Reflective Equilibrium", 29.

<sup>85</sup> Daniels, "Wide Reflective Equilibrium", 29-30.

<sup>86</sup> Daniels, "Wide Reflective Equilibrium", 30.

<sup>87</sup> Daniels, "Wide Reflective Equilibrium", 31.

Daniel writes, is that we indeed need to give an answer about liability of moral judgment, but that it does not show that considered judgments are implausible when they lack initial credibility.<sup>88</sup> Another reason that critics give against the credibility of considered judgments, is the existence of variation and disagreement. However, this cannot count as an argument against them, Daniel says. Even though there is widespread *disagreement*, there is also a lot of *agreement* on many of the considered judgments. Denying credibility on the basis of disagreement would therefore be shortsighted.<sup>89</sup>

Nonetheless, this argument does not give us a solution for the problem either. Indeed, we should acknowledge the fact that there is agreement on a wide spread of moral judgments, but this agreement still does not tell us *why* we are justified believing in a considered judgment. The agreement that exist does not eliminate the possibility that the outcome of the method of reflective equilibrium is completely *unreasonable*.

This means that Daniels' arguments against the problem stated in this paper do not offer a solution. However, we might find an answer to our problem in the arguments presented by Thomas Scanlon.

### 4.3 Expanding the Method

In 'Lecture 4' of his book "*Being realistic about Reasons*", called "*Epistemology and Determinateness*", Thomas Scanlon focusses on an account that shows how we can arrive at knowledge about a subject by thinking of that subject in the right way.<sup>90</sup> The 'subject' here is a normative truth.<sup>91</sup> However, he writes, if the account offered for knowledge about this is in itself a normative claim, how, then, do we know that *this* claim is true? According to Scanlon, this is done with the method of reflective equilibrium.<sup>92</sup>

The understanding and the use of the method is the same one as we discuss in this paper: the method of reflective equilibrium that Scanlon is talking about involves the same "back and forthgoing"<sup>93</sup> In Rawls's theory, the account for *normative truths* (being principles of justice) is the original position, and the justification for those principles is given through reflective equilibrium.

Scanlon discusses why having a judgment among our considered judgments, makes us justified in accepting this judgment.<sup>94</sup> This point is a very relevant one for this paper, since I argue that we are in fact *not* justified in accepting that unless we adjust the starting points for the equilibrium with the consequence that the method loses its justifying force.

In short, as well as Daniels, Scanlon argues that the problem as stated in this paper comes from a misunderstanding about considered judgments. According to Scanlon, we don't decide whether or not we have good reason to believe in them *before* the method, but this decision is actually *part* of the method of reflective equilibrium. The considered judgment should not only be confidently held, but need to seem reasonable while thinking of it under the right conditions. The possibility of having any implausible implications or presuppositions is ruled out, since being a considered judgment implies not having implausible implications.

---

<sup>88</sup> Daniels, "Wide Reflective Equilibrium", 32.

<sup>89</sup> Daniels, "Wide Reflective Equilibrium", 32.

<sup>90</sup> Scanlon, "Epistemology", 71.

<sup>91</sup> Scanlon, "Epistemology", 70.

<sup>92</sup> Scanlon, "Epistemology", 71.

<sup>93</sup> Scanlon, "Epistemology", 77.

<sup>94</sup> Scanlon, "Epistemology", 79.

The pursuing of the method of reflective equilibrium is how we find that some starting points are more reasonable than others.

The justificatory force, Scanlon writes, lies in how reflective equilibrium is reached.<sup>95</sup> This could be identified if we considered the fact that different people might reach different equilibria – which is, as we have seen, what Rawls considered as well in his *Political Liberalism*. To find out how the equilibrium was reached, one should ask two questions: “Did this person reach different conclusions than I did because she began with different considered judgments than I did? If so, should I accept those different judgments?” And: “Did the person reach a different conclusion because she considered different principles than I did, or made different choices than I did about whether to revise a principle or modify a judgment that conflicts with it? If so, should I have made these different choices as well?”<sup>96</sup>

Positive answers should motivate us to change our beliefs so that they will line up with the beliefs of the considered person, whereas negative answers do not give us any reason to change our thinking.<sup>97</sup> This pluralism does not mean that we have to reject either one of the views, according to Scanlon. It might be that the subject matter can be understood in both of the ways equally well.<sup>98</sup> This corresponds to the view Rawls discusses in *Political Liberalism* as well. However, the question may still be asked how the fact that the judgment is not rejected in this process, can be seen as sufficient justification.<sup>99</sup>

This, Scanlon argues, can be examined by answering the question whether or not a person had good reason to doubt her considered judgments and if she made reasonable decisions in the steps of the process of reflective equilibrium.<sup>100</sup> Whenever we think something is reasonable to believe because this *seems* to be the case, and it *seems* like there is no reason to doubt this belief, we can decide to count this belief among our considered judgments.<sup>101</sup> Scanlon concludes that “the justificatory force of the fact that we have arrived at certain judgments in reflective equilibrium depends on the substantive merits of the judgments we make along the way, in beginning with certain considered judgments.”<sup>102</sup>

Nevertheless, this is looking much like a reformulation of the problem that this paper points out, which is that the considered judgments need some justification *before* the method of reflective equilibrium in order to have a justified outcome of that method.<sup>103</sup> Scanlon believes this premise is invalid. He points out that a considered judgment should not only be confidently held, but it also needs to be something that seems clearly true when thought of the subject matter under good conditions.<sup>104</sup>

However, this leads back to the concern Kelly and McGrath emphasized in their paper, being that the most interesting part of it all is not the method of reflective equilibrium, but the reasons why some starting points are reasonable.<sup>105</sup> This concern however, Scanlon writes, is based on a misconception. He points out that the decision to treat something as a

---

<sup>95</sup> Scanlon, “Epistemology”, 79.

<sup>96</sup> Scanlon, “Epistemology”, 79.

<sup>97</sup> Scanlon, “Epistemology”, 79.

<sup>98</sup> Scanlon, “Epistemology”, 80.

<sup>99</sup> Scanlon, “Epistemology”, 80.

<sup>100</sup> Scanlon, “Epistemology”, 81.

<sup>101</sup> Scanlon, “Epistemology”, 81.

<sup>102</sup> Scanlon, “Epistemology”, 82.

<sup>103</sup> Scanlon, “Epistemology”, 82.

<sup>104</sup> Scanlon, “Epistemology”, 82.

<sup>105</sup> Kelly and McGrath, “Is Reflective Equilibrium Enough?": 353.

considered judgment is part of the method itself.<sup>106</sup> One should ask whether the judgment has any implausible implications or presuppositions. Treating it as a *considered* judgment implies that it does not.<sup>107</sup> According to Scanlon, the process of pursuing reflective equilibrium is in fact the way we comprehend why some starting points are reasonable.<sup>108</sup>

Even though his arguments offer us some more understanding of what it is to count something as part of our considered judgments, also Scanlon does not offer a solving account for the problem of the method here. The fact that we need to think of the considered judgment *under the right conditions* is itself an adjustment of the starting points. Even though this adjustment may indeed lead to a just outcome of the method, the justification is again not given *by* the method itself.

Scanlon will probably argue that this is a misunderstanding as well, since thinking of that considered judgment *under the right conditions* is already part of the method of reflective equilibrium. However, by doing this, Scanlon expands the scope of reflective equilibrium. Instead of only justifying the principles that follow from the considered judgments, the method will also answer question as to whether or not a belief might count as a considered judgment. This, in turn, adds another problem to the method. By expanding the method, reflective equilibrium loses not only the justificatory power, but any power it originally had such as the guarantee for coherency in the outcome of the method.<sup>109</sup>

By expanding the method further and further, we will eventually end up with what we started with: ‘regular’ deliberation about moral questions. Expanding the method makes reflective equilibrium unpractical and reduces it to ‘day to day’ moral reasoning.<sup>110</sup> On top of that, I do not see how adding the questions about considered judgments in the method itself, answers the question why considered judgments are justified. Given the method as presented, I think it will *still* just produce coherency. In order to have a just outcome, then, the judgments we reason about *in* reflective equilibrium to find out whether or not to include them in our considered judgments, *still* need to be justified themselves before being brought into the method.

## 5. Justification through Coherency: A fundamental Problem

Thus, the problem this paper started with remains unsolved. The method Rawls introduced in his *Theory of Justice* leads to the question *what* makes considered judgments justified. The method of reflective equilibrium seems to guarantee only coherency: and that does not imply justification. As we have seen, the problem of justification of considered judgments is *not* solved by Scanlon or Daniels.

Indeed, Rawls could and would possibly argue that the considered judgments *are* in fact justified, because they are the considered judgments of *rational and reasonable* people, and that therefore coherency *does* suffice for justification. To say that the considered judgments are the judgments of *rational* people, is to edit the starting point. Still, that way the justification is *not* given by the method itself.

A possible outcome for Rawls were the adjustments of his *Justice as Fairness* that he presented in his works *Political not Metaphysical* and *Political Liberalism*. However, the

---

<sup>106</sup> Scanlon, “Epistemology”, 84.

<sup>107</sup> Scanlon, “Epistemology”, 84.

<sup>108</sup> Scanlon, “Epistemology”, 84.

<sup>109</sup> De Maagt, “Reflective Equilibrium and Moral Objectivity”, 460–462.

<sup>110</sup> De Maagt, “Reflective Equilibrium and Moral Objectivity”, 462.

'politicizing' of his theory *only* adjusts the starting points of the method of reflective equilibrium. Even though Rawls accepts the existence of different reflective *equilibria* and reasonable pluralism, these changes only affect the *input* of the method: not the method itself.

Defenders of the method and indeed Rawls himself might argue that the adjustments made do lead to a *justified outcome* of reflective equilibrium. However, this justification still rests on the adjustments of the starting points, rather than on the method itself. On top of that, while making the adjustments to the starting points so that they will lead to a justified output of the method, several assumptions need to be made as to what can count as just and what justice would be. This way, the theory loses the neutral character it claims to have: it will be biased.

The method of reflective equilibrium, the goal (to produce coherency) and the function, to justify principles of justice, remains the same: including the problem that this method *does not* justify. On top of this, the theory will be biased. This makes the whole justifying process altogether doubtful.

## 6. Conclusion

In this paper, I argued that the method of reflective equilibrium as described by Rawls is problematic, because it cannot provide for the justification as it is claimed it does. I showed that the purpose of reflective equilibrium is to produce coherency. However, for this coherence to be justified, one needs to adjust the input of the method. This is what the problem boils down to: if one is to adjust the input of the method so that it is *already* justified, the justificatory force of the method gets lost.

I showed that Rawls in fact adjusts the conditions for what is an acceptable starting point in his 'Political Liberalism'. First of all, his conception of justice should be understood as a *political* conception. Second, the parties are reasonable and rational people, with two moral capacities: a sense of justice and a conception of the good. Reflective Equilibrium needs to contain at least an account of reasonable pluralism and a distinction between political and non-political beliefs, so that priority can be given to the first kind of beliefs.

Even though these changes make Rawls's theory more realistic, they do not give the method of reflective equilibrium the justificatory force it was missing before. Rawls could argue that with his changes, the produced coherency of the method would in fact be justified. Admittedly, the way Rawls modified his starting points can in fact influence the outcome so that the principles will be just. However, this is the case because he justified the starting points *before* used as input for the method. That way, reflective equilibrium still does not have the justificatory force it is said to have.

I discussed defenses written by Norman Daniels and Thomas Scanlon, who claim that the critique that the coherency cannot account for justification is based on a misconception, especially on a misconception about the considered judgments. We should not compare moral judgments with observation reports, according to Daniels. We do not have the ability to add initial credibility to moral judgments, and even though we do need *some* account about this, it does not make considered judgments completely implausible. Scanlon expands the scope of the method, claiming that the reasonableness (and therefore justification) of considered judgments is part of the method of reflective equilibrium. Nevertheless, their answers still don't solve our problem. If we can't add initial credibility to our considered

judgments, then it will even be harder to examine *why* they are in fact justified for other reasons than reflective equilibrium. Furthermore, expanding the scope of the method makes the method not only unpractical, it also does not give us a solution. Judgments that *can* count as considered judgments will *still* need justification *before* the method.

The problem is that the method of reflective equilibrium cannot justify principles. If we want a justified outcome of the method, we need to make sure the input is already justified in one way or another. This might indeed be a solution for Rawls's theory in general: if we all can agree on a certain set of acceptable conditions for starting points, at least the theory will generate just principles: this might be an interesting inquiry as well. However, these justified principles are not the product of the method of reflective equilibrium as Rawls had in mind. In other words, the justification of the principles can be settled by adjusting the starting points of reflective equilibrium, but this undermines the idea of the theory: to generate just principles through exercising the method of reflective equilibrium. In adjusting the starting points in such a way that the judgments are *already* justified, the theory also has the risk to get biased. This in turn will make the theory lose its whole neutral character.

The theory being biased because it needs to make assumptions about justice for the justification of starting points is one possible consequence of the lack of justifying force of reflective equilibrium. I think it would be interesting to examine how far the consequences of this lack work through Rawls's whole theory. To see which parts of *Justice as Fairness* are affected by the problem reflective equilibrium has, might lead to a better understanding and/or new insights about theories of justice. However, in this paper I did not have the space for that.

The aim of Rawls was, indeed, very humble. To actually produce a theory that would be neutral with regards to conceptions of justice, that will produce justified principles with help of the method of reflective equilibrium, would be groundbreaking. However, as we have seen, this is not the case. The theory Rawls presented lacks justifying force and is, therefore, in need of biased starting points for, ironically enough, acceptable starting points.

## Bibliography

- Daniels, Norman. "Reflective Equilibrium and Justice as Political". In *Justice and Justification: Reflective Equilibrium in Theory and Practice*. (Cambridge: Cambridge University Press, 1996). 144-176.
- Daniels, Norman. "Wide Reflective Equilibrium and Theory Acceptance in Ethics". In *Justice and Justification: Reflective Equilibrium in Theory and Practice*. (Cambridge: Cambridge University Press, 1996). 21-46.
- Daniels, Norman. "Reflective Equilibrium". In *The Stanford Encyclopedia of Philosophy* (fall 2018 edition). ed. Edward N. Zalta. <https://plato.stanford.edu/archives/fall2018/entries/reflective-equilibrium/> (accessed September 11th, 2018).
- Kelly, Thomas and McGrath, Sarah. "Is Reflective Equilibrium Enough?". In *Philosophical Perspectives* 24 (1): 325-359.
- De Maagt, Sem. "Reflective Equilibrium and Moral Objectivity". *Inquiry: An Interdisciplinary Journal of Philosophy*. no. 5, (2017): 443-465.
- O'Neill, Onora. "The Method of a Theory of Justice". In *John Rawls, A Theory of Justice*. (Leiden: BRILL, 2013). 21-35.
- Rawls, John. *A Theory of Justice*. Revised Ed. (Cambridge: Harvard University Press, 1999).
- Rawls, John. "Justice as Fairness: Political not Metaphysical". In *Communitarianism and Individualism*. Ed. Shlomo Avineri and Avner deShalit. (Oxford: Oxford University Press, 1992). 186-204.
- Scanlon, Thomas M. "Epistemology and Determinateness". In *Being Realistic about Reasons*. (Oxford: Oxford University Press, 2014). 69-104.
- Scanlon, Thomas M. "Rawls on Justification". In *The Cambridge Companion to Rawls*. Ed. Samuel Freeman. (Cambridge: Cambridge University Press, 2002). 139-167.
- Wenar, Leif. "John Rawls". In *The Stanford Encyclopedia of Philosophy* (spring 2017 edition). ed. Edward N. Zalta. <https://plato.stanford.edu/archives/spr2017/entries/rawls/> (accessed June 20<sup>th</sup>, 2018).