

UNIVERSITY OF GOTHENBURG

DEPARTMENT OF PHILOSOPHY, LINGUISTICS AND THEORY OF SCIENCE

## **Is Fully Informed Better Than Cast in Oblivion?**

- A discussion on The Veil of John Rawls and The Proviso of David Gauthier -

Bachelors Thesis in Practical Philosophy

October, 2009

Author: Tomas Halberstad

Advisor: Christian Munthe

<b>1. Introduction</b>	<b>3</b>
1.1. Background	3
1.2. Purpose and Problem	3
<b>2. The Veil of Ignorance</b>	<b>4</b>
2.1. Restrictions of and Reasons for the Veil	4
<b>3. Gauthier's Ideas and Point of View</b>	<b>5</b>
3.1. Gauthier's Theory – Concepts One Through Three	6
3.1.1. Rational Bargaining Between Jay and Dee	9
3.2. Comparative Intermission: The Purpose and Nature of the Contract	10
3.3. Gauthier's Theory – Concepts Four and Five	11
3.4. Timeline Correction	12
3.5. Comparative Intermission: Pre-Theoretical Freedom and Archimedean Point	13
<b>4. Does Gauthier succeed?</b>	<b>13</b>
4.1. Lifting the Proviso	13
4.1.1. The Guard and the Swap-Meet	14
<b>5. Can we use the theory anyway?</b>	<b>16</b>
5.1. The Difference in Normative Content	17
5.2. Choice From Distorted Judgements	18
5.3. Criteria of Impartiality	19
5.4. Does Fairness Imply Strong Impartiality?	20
<b>6. Summary</b>	<b>21</b>
6.1 Conclusion	22
<b>Bibliography</b>	<b>23</b>
Other Resources	23

# 1. Introduction

In his book, *A Theory of Justice*, John Rawls put forward ideas that blew life in the then dormant debate on Social Contract Theory.<sup>1</sup> Those ideas raised a lot of questions. One, brought forward and discussed in this essay originates from the intuitive idea that full knowledge must be better than restricted knowledge when it comes to making some kind of decision, or choice, on a subject or matter. The idea is that it is better to be fully informed than cast in oblivion.

## 1.1. Background

The questions concerns Rawls' initial part to his contracting situation, *The Original Position*, and in particular the concept of *The Veil of Ignorance*.<sup>2</sup> This Veil restricts the knowledge of the individuals in the contracting situation. Rawls introduces the concept in order to be able to reach a determinate, and by him preferred, result. But does he need to restrict knowledge to reach this preferred result? Is there another theory or idea which arrive at a similar conclusion, or a different one but equally valid, without imposing these, or any, knowledge restrictions?

In the book *Morals by Agreement*, David Gauthier is arguing for an initial part, of a contract situation, called *The Initial Bargaining Situation*.<sup>3</sup> The main difference to Rawls' initial part being that the individuals are fully informed; no knowledge restrictions are imposed on them.

Still, Rawls has reasons for introducing The Veil (as I will henceforth call it). These, to be brief, are about preventing the individuals from acting in an undesirable manner. In effect, Gauthier has to be asked: What (in his theory) prevents the individuals in the contracting situation from acting in an undesirable manner? If there is *nothing*, how does these individuals arrive at an agreement (one argument of Rawls' is that without The Veil, and its knowledge restrictions, an agreement can probably not be reached)? If there is *something*, will it do to bring forward an outcome deserving to be labelled just or, at least, morally justified?

## 1.2. Purpose and Problem

The purpose of this essay is, therefore, to examine Gauthier's critique of Rawls' Original Position as presented in *Morals by Agreement*. The particular object of criticism is Rawls' concept The Veil of Ignorance and the restrictions on knowledge that it brings to, or puts on, the individuals in a hypothetical contracting situation, supposed to determine the just

---

<sup>1</sup> Rawls, John. 1971.

<sup>2</sup> For a full account of The Original Position and The Veil of Ignorance see Rawls, John. 1971, p. 118f.

<sup>3</sup> Gauthier, David. 1986.

distribution of liberty, opportunity and material goods. Gauthier puts forward a theory where no restrictions on the knowledge of the individuals in the contracting situation are needed. The main question to answer is: Does Gauthier's Initial Bargaining Position provide a better basis for a social contract theory than Rawls' Original Position, and thereby cater to the intuitive idea stating that fully informed is better than cast in oblivion, when it comes to making a choice (of this kind)?

## 2. The Veil of Ignorance

Rawls' aim in designing The Original Position is to design an initial bargaining situation that would not only seem, but also be, fair among all individuals underwriting or agreeing to the hypothetical social contract. These individuals are assumed to be maximizers of their own self-interests as well as mutually disinterested.

The idea is that if all information, of relevant nature, is taken into account and all individuals are fairly situated, then the principles agreed upon in the contracting situation would also be fair. The fairness would be transferred from the initial bargaining position to whatever is agreed upon, and the agreement will thus describe a justified claim about the traits of a just society. In other words: Rawls needs to situate the individuals in a condition that ensures fairness, for them to reach a fair agreement. This is where The Veil comes in.

### 2.1. Restrictions of and Reasons for the Veil

Rawls introduces The Veil partly for the reason that without the knowledge restriction that it brings, bargaining in the contract situation would, he suggests, be hopelessly complicated.<sup>4</sup> More important, however, is a pre-text to the idea in which Rawls wants to make extinct the effects of specific coincidences, or matters of chance.<sup>5</sup> This because the individuals should not be able to use, or exploit, individually favourable circumstances to their own advantage and in that causing disharmony in the contracting situation. What Rawls want is a simple and harmonious bargaining process and he also wants a determinate and fair result from it. He sees that his restrictions on particular information in the contracting situation are necessary to achieve this.<sup>6</sup>

---

<sup>4</sup> Rawls, John. 1971, p. 140.

<sup>5</sup> Ibid. p. 136f.

<sup>6</sup> He is not happy with just any determinate result. He explicitly says that he wants do define the contracting situation so that the preferred outcome, on his view the *Principles of Justice* (more commonly known as *The Liberty Principle* and *The Difference Principle*), is reached.

The knowledge restricted by The Veil in order to get determinate results are in general, and as Rawls sees it, knowledge that can distort the judgement of the choosing parties. The idea here being that the result of an agreement, if one could be reached, between individuals with distorted judgements would be unfair.<sup>7</sup> The individuals, in the contracting situation, are left knowing the general facts about the world, the society and the institutions but are in oblivion regarding the particulars of themselves. Rawls assumes that it stands clear that since the individuals don't know anything about their differences in opinion and that everybody is equally rational, and situated similarly, then everybody is persuaded by the same arguments, and if everybody is persuaded by the same arguments, a unanimous agreement *can* be reached.<sup>8</sup>

A related reason for The Veil is to let the individuals be able to produce an unbiased assessment.<sup>9</sup> Rawls has an idea that theories of justice, in particular, should be able to live up to fundamental conditions of impartiality and The Veil can be seen as a means for that end as there, under it, is no room for biased behaviour on any part.<sup>10</sup> Thus, The Veil is to be seen as a part of his main objective with the contracting situation which is, as stated, to design a situation where rational, self-interested parties will reach an unanimous agreement. Furthermore he also wants the principles chosen in the contracting situation to be the result of a decision procedure that is structured in such a manner that allows it to incorporate pure procedural justice at the highest level. In order for the procedure to be structured in such a manner, Rawls argues that The Veil is a necessity.<sup>11</sup> Presumably, it is the fairness of the agreement produced by The Veil that explains this necessity.

### **3. Gauthier's Ideas and Point of View**

Can a determinate result, and a result like the one Rawls envisions be reached without The Veil? Gauthier thinks that it can, maybe not in regard to the principles that Rawls want to be chosen but in regard to the moral quality (the justice) of the society that Rawls expect his principles to lead to. Gauthier thinks that his individuals seem "...well suited to join the kind of

---

<sup>7</sup> Freeman, Samuel. 2009.

<sup>8</sup> Rawls, John. 1971, p. 84, 358f.

<sup>9</sup> In Rawls' case "...of the justice of existing social and political institutions and of existing preferences and conceptions of the good." Freeman, Samuel. 2009.

<sup>10</sup> When I write *Impartiality* I mean *Moral Impartiality* as distinguished in Jollimore, Troy. 2008.

<sup>11</sup> Rawls, John. 1971, p. 139.

society that Rawls [envisions]”.<sup>12</sup> Add to this Gauthier’s thought that the basic idea of The Veil is wrong and the stage is set for this essays main question (as stated in 1.2.).

Gauthier’s primary reason for concern is not, however, that The Veil restricts knowledge but that the introduction of it is motivated by moral intuition. This, Gauthier argues, means that the arguments for Rawls’ entire theory *not only* proves that choices made in the contract situation are rational, but also that substantial moral premises are presumed as given. Rawls does not deny this but argues that this is the way it has to be done if, one: the theory as a whole is to be coherent, and two: there is to be a theory of justice at all!

Gauthier’s goal at a general level is, if not to all in all refute the *coherentism* of Rawls, then at least to prove, or at a minimum put forward a plausible idea, of *foundationalism* as far as moral epistemology is concerned. He is interested in resolving the, on his view, foundational crisis of morality and thinks that this quest will find a “... temporary harbour in ... [his] book.”<sup>13</sup>

### 3.1. Gauthier’s Theory – Concepts One Through Three

Gauthier wants to use contract theory thought to derive morality from rationality (and this without adding morality, or any kind of moral norm, as a premise to rational individuals in the contracting situation, as one might claim that Rawls does).<sup>14</sup> His main purpose is, then, to answer the moral sceptic and find an answer to the sceptic’s question: Why be moral?

The chain by which he works is to first connect interest or self-interest, individual utility, subjective utility, to a form of practical reason.<sup>15</sup> Then he connects practical reason with reason, and then that idea of reason connects with morality.

An aim with the theory is to create harmony between individuals, or more to the point: harmony between the interests of self-interested individuals. This created harmony is, in Gauthier’s terminology, an *artificial harmony*. He assumes (like most, if not all, other social contract theorists) that there can be no *natural harmony* between individuals. This, since the self-interests of individual’s conflict. However, there is much to gain for each individual able to co-operate. A state of natural harmony, albeit idealized and abstract, must, however, be established as both starting point and point of aim (or outline) for the theory to be. This state,

---

<sup>12</sup> Cudd, Ann. 2008.

<sup>13</sup> Gauthier, David. 1986, Preface.

<sup>14</sup> Gauthier actually works in the contractarian tradition, Rawls in the Contractualist. To distinguish in this manner is of no need to me, therefore I refer both of them to contract theory thought. For an account of Contractarian and Contractualist thought see D’Agostino, Fred. and Gerald, Gaus. 2008.

<sup>15</sup> When Gauthier writes *Utility* he means “objects of preference relation”. He also “... identifies rationality with the maximization of utility”. Gauthier, David. 1986, p. 22.

of natural harmony, is according to Gauthier, a point that later can be used as "... a foil against which morality appears more clearly."<sup>16</sup> This idea, this idealisation, *The Morally free Zone* is a version of the perfectly competitive market and it is the first of five concepts central to Gauthier's theory. This is a market where there is no tension between the maximization of the social and the individual good. Thus, individual maximization will also maximize the room for mutually beneficial co-operative ventures.

However, according to Gauthier, there can't be such a "perfectly competitive market" between individuals unless these are assumed to meet constraints that real people seldom live up to. For a less idealized and abstract model of a well-functioning competitive market, moral constraints, of sorts, on the individuals and their pursuit of self-interest is required. This because:

"... co-operative activities almost inevitably involve a prisoner's dilemma: a situation in which the best individual outcomes can be had by those who cheat on the agreement while the others keep their part of the bargain. This leads to the socially and individually sub-optimal outcome wherein each can expect to be cheated by the other. But by disposing themselves to act according to the requirements of morality whenever others are also so disposed, they can gain each others' trust and co-operate successfully."<sup>17</sup>

Rawls is a Coherentist. He wants his theory to work, as a whole; he wants it to be as coherent as possible and though Gauthier also wants, and claims, his theory to be coherent his main concern is with founding morality in rationality. This is important for Gauthier not only for the success of his theory but also for giving a possible solution to the compliance problem: "The problem of justifying rational compliance with the norms that have been accepted... [and solving this problem] ... must [according to Gauthier] always drive the justification of the initial situation and the conduct of the contracting situation."<sup>18</sup> With all this in mind Gauthier begins to model his contracting situation and the individuals therein.

To begin with, Gauthier introduces his second concept: *The Principle of Minimax Relative Concession*. Gauthier looks at the contracting situation as a bargain and in this bargain the individuals are, for themselves, wanting, and trying to, negotiate and agree to (moral) rules, which will maximize their own (expected) utility. Now, when entering this bargain the individuals already have different amounts of utility, which means that they all have an

---

<sup>16</sup> Gauthier, David. 1986, p. 13

<sup>17</sup> Cudd, Ann, 2008.

<sup>18</sup> Ibid.

individually unique stake in the bargain. So, The Principle of Minimax Relative Concession is linked to the equal rationality of the individuals. According to Gauthier, the most rational outcome to choose, when choosing between different expected outcomes, for ones initial stake is "... that the greatest [possible] concession, measured as a proportion of the conceiver's stake be as small as possible."<sup>19</sup> There is also an equivalent in *The Principle of Maximin Relative Benefit* which means that "... the least relative benefit ... be as great as possible."<sup>20</sup> In other words: The most reasonable outcome for each individual, as far as maximizing utility is concerned, is that they see to it that they minimize there maximum relative concession and maximize there minimum relative benefit, and do so, as the word relative points out, in regards to each of the other individuals in the bargain. If the concession and benefit will be deemed reasonable then agreement can be reached. Gauthier also claims that this concept; these two principles, "... captures the idea of fairness and impartiality in a bargaining situation, and so serves as the basis of justice."<sup>21</sup> (For short the two principles will now be referred to as *Minimax* and *Maximin*.)

For these principles to work a concept of rationality, and equal rationality, must be put forward, and so Gauthier does. He calls this concept, his third, *Constrained Maximization*. What he means by this is that to be rational is to be "... disposed to comply with mutually advantageous moral constraints, provided [one] expects similar compliance from others."<sup>22</sup> When one so does one is a constrained maximizer. This can be set against a person who Gauthier might call a *Maximizer* or a *Maximizer Period* who is disposed only to maximize utility, at all times regardless of what everyone else is disposed to comply with.<sup>23</sup>

These three concepts constitute, on my view, the actual bargaining situation, as put forward by Gauthier. We have The Morally Free Zone; the perfectly competitive market, which, with its natural harmony is the template from which the artificial harmony of co-operating, self-interested individuals will rise. There is also the two principles of Minimax and Maximin. They are there to secure that each individual concedes as little as possible, and also benefits as much as possible, from the agreement and in proportion to ones stake. Last we have the

---

<sup>19</sup> Gauthier, David. 1986, p. 14.

<sup>20</sup> Ibid.

<sup>21</sup> Ibid.

<sup>22</sup> Ibid. p. 15.

<sup>23</sup> Gauthier assumes that constrained maximization is the most rational way to be if rationality, or to be rational, is to maximize. This because he also, at first, assumes that the "net advantage that constrained maximizers reap from co-operation exceeds the exploitative benefits that others may expect". He also goes to some length and through many technicalities to prove this. If his assumptions are right or wrong are, however, for the purpose of this essay irrelevant.

concept of Constrained Maximization which says that rationality is to maximize and to maximize in the best way; to gain the most, one puts moral constraints on ones pursuit of utility in particular situations.

### 3.1.1. Rational Bargaining Between Jay and Dee<sup>24</sup>

The following is a simplified example of why Gauthier thinks Minimax is needed and how it will work in general: Jay and Dee are equally rational. They live in the world of Pre-Bargaining. This is not a very nice world. It is a world where everyone is working individually (for themselves, by themselves) to gain utility.

Because of Jay's and Dee's equal rationality they see that if they were to co-operate, their work would amount to more utility, even more than their present, individually gained, utility put together. They both would throw themselves at the chance to co-operate if it weren't for one small problem: How do they divide the, jointly gained, utility?

Here is why this is a problem: In the world of Pre-Bargaining Jay has been working rather hard whilst Dee only has worked moderately hard. This means that the individual amounts of utility that Jay and Dee bring into the bargain are different.

For the sake of keeping this example simple we will now assume that the different amounts of utility of Jay and Dee are linear to monetary values.<sup>25</sup> In this context we would now say that Jay has more money than Dee. And since he does, Jay has a higher stake in the bargain.

Jay looks at how much money he would receive if he were to receive all the money from co-operation. (This after covering Dee's costs.) Dee does the same. Jay arrive at a figure of 500 Swedish Krona (SEK) and Dee at the sum of 50 SEK. Then both, individually, compare that sum to the sum they would have got if there were no co-operation. Again to keep it simple, let us say that there would be a 0 SEK pay-off on their individual stake if there were no co-operation. The respective differences between the two sums, for Jay 0-500 SEK and for Dee 0-50 SEK, are their concession. These numbers are needed for Jay and Dee in order to find out what their *relative* concessions are, or can be. This since none of the two should gain more, or loose more, at least when put in ratio, than the other (for short: to ensure that there is a, fair, balance in the distribution of money).

Jay and Dee do a little calculating and soon arrive at a relative concession (a ratio) of 0.294. This means that Jay should accept the outcome of receiving 353 SEK ( $500 \times 0.294 = 147$ ,  $500 - 147 = 353$ ) through co-operation and Dee should accept an outcome of  $\approx 35$  SEK.

---

<sup>24</sup> An example in part borrowed from Gauthier, David. 1986, p. 137f

<sup>25</sup> As Gauthier does in one of his examples. Ibid.

This because any other ratio, than 0.294, would require one of them to make a greater relative concession, hence making the agreement unfair.<sup>26</sup>

### 3.2. Comparative Intermission: The Purpose and Nature of the Contract

Let's pause here to see where Rawls stand on the Gauthierian concepts discussed so far, to get an overview of, and show, where the real conflict, for the purpose of this essay, lies.<sup>27</sup> Rawls has a different goal than Gauthier. Rawls wants his individuals to choose his principles of justice (instead of Gauthier's principles of Minimax and Maximin). Then, from them, choose a society. However, Gauthier sees his theory to "... produce liberal individuals, who seem well suited to join the kind of society that Rawls [envisioned]."<sup>28</sup>

As now touched upon, there is conflict between the two philosophers in the matter of which principles are to be chosen, but since this essay is concerned, roughly, with the idea of individuals that are fully informed versus those that are cast in oblivion, and this prior to choice or regardless of what choice is made, this conflict is disregarded for now.

Then we have rationality. On Rawls' view, to be rational is to choose the most efficient means to an end and since Rawls' individuals also are self-interested, the end, more often than not, is to benefit oneself. Gauthier might take it a step further but only because he has to. Since he has no Veil of Ignorance the idea must go a little deeper than to just want to maximize. To describe the possibility of an agreement or contract, Gauthier invokes the Minimax, Maximin and Constrained Maximization notions. Still I do not see how Rawls would have a problem with this. Rawls says, "... as far as possible the concept of rationality should be kept as narrow as possible."<sup>29</sup> How could he then have a problem with a narrower version of the same concept? Likewise I fail to see why Gauthier would have any objection to Rawls' assumption that to be rational also includes to not being, feeling or acting envious. In fact, Gauthier makes a similar claim when he excludes envy because he assumes it has a quality of being a *relative preference*.<sup>30</sup> An assumption that on Gauthier's part only streamlines the theory further (as it does with Rawls' theory).

---

<sup>26</sup> Gauthier puts forward arguments as to why such an agreement should be accepted by both and why it is fair. He also acknowledges that minimax relative concession does not always result in equal relative concessions. Both arguments and acknowledgements are, however, of lesser importance to this essay and therefore left out.

<sup>27</sup> I use the term *Conflict* in a broad sense meaning some thing like "not saying the same, or, having different meaning. The term, and the opposite term *Non-Conflict*, is used in this way throughout the essay.

<sup>28</sup> Cudd, Ann. 2008.

<sup>29</sup> Rawls, John. 1971, p. 14.

<sup>30</sup> Heath, Joseph. 2009.

### 3.3. Gauthier's Theory – Concepts Four and Five

If concepts one through three constitute the actual bargaining situation, the actual contract situation, why does Gauthier need two more concepts to complete his theory? He seems to have already established a goal, the disposition of the individuals and two principles of co-operation. This, however, is only the procedure that, on Gauthier's view, leads to the solution. Of equal importance is the starting point, a point from which negotiations begin, pre-procedure if you will. This pre-procedure defines what stakes can be legitimately brought into the bargain. That is, not just any stakes are allowed when entering the contract agreement. This is where Gauthier's fourth concept, *The Proviso*, comes in.<sup>31</sup> It can be seen as a rule of conduct regarding the accumulation of stakes to take into the bargain which "... prohibits bettering one's position through interaction worsening the position of another."<sup>32</sup>

Gauthier describes the rationale of this part of the theory thus: "If persons are willingly to comply with the agreement that determines what each takes from the bargaining table, then they must find initially acceptable what each brings to the table."<sup>33</sup> If it isn't so, if the persons initially do not accept what each brings to the table then no deal, no agreement can be struck. The idea here being that the initial position, the state *before* bargaining has begun but *after* it has been decided that bargaining *is* to take place, must; by necessity, have been reached by the individuals without outer force, non-coercively. "... if what some bring to the table includes the fruits of prior interaction forced on their fellows, the initial acceptability will be lacking."<sup>34</sup> No one wants to *first* make a concession then *later* be the beneficiary of something that they already feel belong to him or her. In short, The Proviso assures a condition of what we may call pre-theoretical freedom; no one has the stake he has in the bargain due to the use of force or manipulation by, or on, others.

The fifth, and last concept, to Gauthier's theory is that of *The Archimedean Point*, which, on Gauthier's view, is a hypothetical vantage point from which an observer (presumably an ideal observer or in Gauthier's terminology, an *Ideal Actor*) can objectively perceive the theory as a whole, and do so with a view of totality.<sup>35</sup> Gauthier claims that at an Archimedean point

---

<sup>31</sup> Gauthier initially uses the term *The Proviso* and does so, with some exceptions, throughout his main discussion, he is though clear, in footnotes and more that he means *The Lockean Proviso* and that this is not an original idea but one derived from Locke, John. 1690, *Two Treatises of Government*, London: Awnsham Churchill. And then later branded *The Lockean Proviso*, and used by Robert Nozic, 1974, *Anarchy, State and Utopia*, New York: Basic Books, pp. 175-82.

<sup>32</sup> Gauthier, David. 1986, p. 16.

<sup>33</sup> Ibid. p. 15.

<sup>34</sup> Ibid.

<sup>35</sup> Archimedean point, *Wikipedia, The Free Encyclopedia*, 2009.

an individual can move the moral world. He writes that “ [it is a] ... position one must occupy, if one’s own decisions are to possess the moral force needed to govern the moral realm.”<sup>36</sup> He also assesses that this point “... must be one of assured impartiality.”<sup>37</sup> This, however, is all a bit hard to grasp and make sense of. Still, Gauthier uses these ideas and notions to evaluate the outcome of his theory, to judge the “... structure of society [and] the basic principles that underlie social interaction.”<sup>38</sup> In short: The Ideal Actor sits on the Archimedean Point. From there The Actor views the four other concepts, the concepts preceding this one. He evaluates them one by one (and also, through Gauthier, compares them to other concepts, and uses of an Archimedean Point brought forward by other thinkers; Rawls and John Harsanyi) then relates them to each-other, then, either directly or indirectly, to this one, the Archimedean Point, in what Gauthier calls *Archimedean Choice*.<sup>39</sup> “In embracing these other concepts central to our theory, the Archimedean Point reveals the coherence of morals by agreement.”<sup>40</sup> One can say that The Archimedean Point is the glue that holds this framework of thought together and (according to him) the impartial choice made therein proves that Morals by Agreement bring forward the best principles to structure a society.

### 3.4. Timeline Correction

I have now presented the five concepts in the order in which they first appear, and are presented by Gauthier. This order is, however, not the order in which events unfold in the contracting situation. So to further explain, make clear and, in short, round up how Gauthier sees the progression, I sum up:

For starters, we have the initial situation where everyone who’s there *has not* bettered their situation through forcing or manipulating interaction worsening the position of another; thus respecting, and fulfilling, the proviso. Or, at least, none of the goods so acquired are brought into the initial situation (and can thus not be enjoyed in the resulting society). Then we have the goal of artificial harmony modelled on the natural harmony of the perfectly competitive market. This goal will be reached through the two principles of minimax and maximin and to reach these principals the individuals are rational and to be rational is to be a constrained maximizer; to realise that one needs moral constraints on one’s pursuit of utility

---

<sup>36</sup> Gauthier, David. 1986, p. 233.

<sup>37</sup> Ibid. p. 16.

<sup>38</sup> Ibid. p. 233.

<sup>39</sup> Ibid. p. 233f.

<sup>40</sup> Ibid. pp. 16-17.

in particular situations. All this will then be viewed and deemed coherent from the Archimedean Point.

### **3.5. Comparative Intermission: Pre-Theoretical Freedom and Archimedean Point**

Now, how would Rawls view these two concepts, the Gauthierian concepts four and five, The Proviso and the Archimedean Point? I opt here to bypass the timeline and begin with looking at the latter.

Rawls himself uses an Archimedean Point and he uses it in a similar manner as Gauthier.<sup>41</sup> That will really suffice to consider it a non-conflict. They both try to formulate their theories so that it will be possible, or so the theories will have, points from which the fundamental structures can be evaluated, the coherence can be viewed, and in some manner act, justificatory. And so we move on to the last concept.

By default we are now down to the proviso and it seems that for the purpose of this essay, and to answer the question (as presented in 1.2), here is where we must examine further. The reason for this is that Rawls' Veil seem to have it's equivalent in Gauthier's Proviso in that they have the same job to do in the philosophers respective theories but do the task, relative to each other, in a complete opposite manner regarding the main focus of this essay. Rawls' Veil restricts knowledge while Gauthier's Proviso is a rule, or principle of sorts that restricts what goods can be transferred from the pre-bargaining situation and into society. The outcome of the job they do is different, though. While The Proviso merely ensures that prior use of force or manipulation is not allowed to enhance the individuals unique stake in the bargain, The Veil excludes not only this, but *all* effective partiality thus undercutting the idea of the contract as an agreement resulting from a *bargain*.

## **4. Does Gauthier succeed?**

For this first part, the critical part of this thesis, Rawls is set aside and instead I will try to find out if Gauthier succeeds in terms of his general goal of solving the foundational crises of morality. Does a foundationalist, instead of a coherentist, way of deriving morality from rationality accomplish this objective? The apparent threat is here The Proviso, since it very much looks like a moral constraint on the bargaining producing the contract.

### **4.1. Lifting the Proviso**

What would happen if we lifted The Proviso? Or perhaps more to the point: Why is it there? Gauthier doesn't want to impose any moral norms on the individuals in the contracting

---

<sup>41</sup> Rawls, John. 1971, p 261f.

situation (as already stated: the moral norms should be derived from the rationality of the parties to the contract). However, the constraint that one shall not better one's position through interaction worsening the position of another (i.e. The Proviso), is surely a moral norm!

Now, it could be the case that The Proviso is put in place just to avoid an even more complex bargaining situation, but one that could be theoretically proven, and then, in that, find a foundationalist theory for morality. Gauthier, however, does not do this.

It *is* the case that Gauthier does not intend The Proviso to be in place *during* the bargaining. It is a mere pre-bargaining threshold in place to secure that only the right, the allowed, types of stakes are brought into the bargaining situation. The feeling, however, is that it's lying there lurking throughout the decision process, but to be fair, Gauthier is never explicit on this issue. If he were, and my inkling was true, he would have, one: a theory that works but two: no foundationalist grounding of morality on rationality. This because, as stated, The Proviso then would be a moral norm constraining the parties much like The Veil in Rawls' theory, hence, *would* influence what agreements that could be struck in the bargaining situation.

#### **4.1.1. The Guard and the Swap-Meet**

Does it matter where The Proviso is "placed" when it comes to what influence it has on the bargain? It does not seem to be the case. Here is an example to illustrate what I mean.

Scenario one: Imagine you have a room with only one point of entry, a door. Outside of that door you have a guard, only letting people bring into the room what regulation allows. To keep things simple we add that the guard does not make mistakes and the people that want to get into the room are rational in the Gauthierian sense (as described in 3.1. and 3.2.). Let us say that a big swap-meet is to take place in the room as soon as the guard has done his job, securing that everyone who is inside only holds items that regulation allows. The guard is finished with his job - The swap-meet can begin! The people in the room start trading and bargaining using the items they have brought. They keep at it all day. When night falls everybody's happy. The swap-meet comes to an end and all the people exit the room, passing the guard, carrying different items than the ones they came with.

This scenario could in one sense be seen as how The Proviso is meant to work, according to Gauthier: Regulating what is brought into the bargain but not functioning during the bargain itself. But, is it all an illusion? A different kind of set-up might clarify.

Scenario two: Again at the swap-meet, but this time the guard sits behind a desk *inside* the room. All of the people are walking around, trying to find, and make, a bargain. Deals are

struck, or at least provisionally, because before every deal is sealed it has to get a stamp of approval from the guard and that stamp is received after inspection. Now, since the guard hasn't been standing by the door the people at the swap-meet have been able to bring in a lot of items not allowed by regulation. All of these items will upon inspection of the provisional deals be discarded and no deal involving such an item will get the guard's stamp of approval (and in that not be able to go through). If such an item is included, and so discarded, the bargainers will most likely abort the bargain (if there is any items left in it). This because, it in all cases will be a worst deal (than the one provisionally agreed upon) for at least one of them.

There might be questions asked here as to why the people in the swap-meet have to come to deals or bargains at all. The answer to that lies in the fact that this examples are parallel to, in general, the whole of Gauthier's Initial Bargaining Situation and, in particular, the concept of The Proviso. If there were no deals struck between the people in the swap-meet there would, in parallel, be no agreement reached by the individuals in The Initial Bargaining Position. And since we are using The Proviso (the guard), in the way that Gauthier intended it, pre-bargaining, this would mean that Gauthier's theory would not come to a determinate result and ultimately fail.

With that in mind and going back to scenario two we can say that there will be deals done (because we are not trying to tear up the theory, just examine how The Proviso really works). Deals only using items that regulation allows. So, at the end of the day the people exit the room, passing the guard behind his desk, carrying different items than the once they came with.

This scenario could *also* in one sense be seen as how The Proviso is meant to work, according to Gauthier: Regulating what is brought into the bargain but not influencing the bargain itself. (It is clear that the guard *does* influence the provisional bargains but maybe not the actual bargains agreed upon; the deals to have gone through at the end of the swap-meet.) This time, however, some might be inclined to say that the guard *does* influence the bargaining (the actual bargains). This, because it intuitively might seem like he does. But, is this also an illusion?

If we compare the major building blocks of scenarios one and two we find only similarities. We have the guard, regulating Items. We have the Items allowed by regulation. and, we have the rational people. Now remember, these people want what is best for themselves, no-matter what state of mind. So, holding a grudge against the guard for discarding items not allowed by regulation will not cloud their judgement towards what is a good deal and what is not a good deal.

This means that the outcome of the two scenarios, the deals agreed upon, will be the same. Even if you defied Gauthier and placed The Proviso post bargaining, the outcome, since there is a demand for satisfying The Proviso, would be the same. Call this, scenario three: Imagine the guard locking the door after everybody had got in and not letting people out until the deals they struck only included items that were allowed by regulation.

The idea here is that Gauthier is trying to get away on a technicality. It is difficult, if at all possible, to show that The Proviso has its hands *in* the agreements that are struck in The Initial Bargaining Position but that it influences what agreements *can* be struck seems crystal clear! In parallel: The guard does not go around telling people how to agree; which bargains to make but he decides what to bargain with or when a bargain is valid. Not convinced?

Lifting The Proviso. Scenario four: No guard. The people can swap whichever items they want and leave when they are done. Will we get a determinate result? Possibly, but remember the rationality of these people. A deal that will make them worse off than they were before entering the swap-meet won't be an option. A determinate result is not a sure thing (and the failure of the whole theory lurks around the corner). However, if a determinate result is reached, will this result, if switched to the idea of the guard as The Proviso, lead to a valid theory of justice? Most certainly not! On this, we have seen, Gauthier would concur.

To conclude: We have not, with this example, proved that The Proviso influences the actual bargain directly. We have, however, proved that The Proviso, and satisfying it, *does* influence what agreements that can be reached. The idea is that with The Proviso in place we have an outcome and no-matter where the proviso is placed in the chain of the Initial Bargaining Position the outcome is the same. This because of one: The allowed stakes and two: The rationality of the people. If The Proviso is lifted the outcome will be a different one. And since the outcome is a result of the agreement we can conclude The Proviso in-directly effects the agreement; The Proviso influences what agreements can be reached. And, since The Proviso is formulated as a moral norm it must be so, that a moral norm influences the agreement.

## **5. Can we use the theory anyway?**

For our project, which, again, is to see if Gauthier can cater to the intuitive idea that fully informed is better than cast in oblivion, the prospect of Gauthier failing with his main project might not be all that damaging. We already accept Rawls' theory and if Gauthier fails in the manner which we have sketched above then Gauthier's theory might still work and work in the same tradition as Rawls. That is, while at the same time deriving morality from rationality, premises of moral norms are imposed upon the rational parties. To clarify: We from now on

assume that Gauthier fails to base morality on rationality alone, since The Proviso is a moral norm that has an influence on what agreements can be reached in the bargaining.

So, the question, as the title of this essay reads, stands. Could we, in some crude manner, remove The Veil from Rawls theory and then substitute it with Gauthier's initial position - thus incorporating fully informed individuals - without thereby damaging its plausibility?

I will try to find this out by comparing two angles of approach using The Proviso instead of The Veil and then trying to assess the outcome. I will also like to point out that the discussion in 5.2. through 5.4. is highly speculative but I do not want to leave any stone unturned before I come to a conclusion.

### 5.1. The Difference in Normative Content

First off, though, we have to say a little more about the difference in normative content behind the two concepts - Justice as Fairness and Morals by Agreement. This because, what they say and through that, how they prescribe, and permit, the individuals to act are first different and second of importance for the outcome. With that said, it should be remembered that neither Justice as Fairness nor Morals by Agreement in general, on The Veil and The Proviso in particular, are theories or concepts about how to control the individuals. They are about how the individuals must be *situated*, and must *be*, for their choices to be normatively valid.<sup>42</sup>

One way to understand the objective of The Veil and the objective of The Proviso finds them similar. This is, however, wrong. The Proviso might say: do not use ill means at your disposal to better your position through interaction worsening the position of another. The Veil might state: you do not have the ill means at your disposal, therefore, you cannot through interaction worsen the position of another. This could, then, be seen as a re-write of The Veil's objective to fit the objective of The Proviso because The Veil actually goes a lot further than the former. In this context The Veil would state: You do not have *any* means at your disposal! This implies that you through interaction *would be allowed to* worsen the position of another if you only had the means at your disposal. It also implies that you, through interaction, would be allowed to worsen the position of yourself. The difference here between The Proviso and The Veil is intent; on the latter you can't have any, since you lack the means to do anything in your self-interest.

For The Proviso this leads to using all the means *allowed* to better your position. Under the Veil you have no means, still, the self-interested individuals want to better their position.

---

<sup>42</sup> Here, and for now, disregarding the greater scope of Rawls theory.

The only way to do this, then, is to see to it that all individuals have as good a position as possible. And, in that securing an as good as position as possible for yourself.

## 5.2. Choice From Distorted Judgements

The reasons for Rawls imposing The Veil on the individuals in the contracting situation have already been stated, (in 2.1.). These reasons will lead to the individuals not acting, or making choices based on, distorted judgements.<sup>43</sup> The objective here being not to let the individuals act in a self-biased fashion.

Now, if The Proviso is to be used in this new way; as a rule of conduct throughout the contracting procedure, and if it is to work, one might ask why Rawls insists on this thick veil as a criteria imposed on the individuals instead of just naming it as The Proviso is: a rule to oblige to?

Both Rawls and Gauthier (through this new viewing of his theory) feel that rationality is not enough to make the individuals choose right and do right (with regards to each other). In a way, and although just an implication, Rawls makes his individuals choose and do right with a setting up of *them*, the individuals, modified through The Veil. Gauthier, however, tries to make his individuals choose and do right with the set up of the *contracting situation*, through The Proviso.

The question then is: If rationality is not enough to make the individuals choose and do right, and that does seem the case or else The Proviso would not be needed, what is it that says that they will not do whatever that they can to try and get around the rule of The Proviso?

This might seem like a non-question. Since (as stated in 5.1.), the interest lies in how the individuals must *be* and not to control what they choose. However, the successes of using The Proviso instead of The Veil (crudely stated) stems from this question.

The idea is this: If The Proviso is used instead of The Veil we inevitably introduce the possibility of choice from distorted (or biased) judgements into the bargaining situation. This can lead to (at least) two things as far as the agreement goes. One: an agreement will not be reached since people prioritise wrong, ultimately failing the whole theory. Two: (if an agreement is reached) the agreement will not be a valid one, or less valid than one using The Veil as far as a theory of justice goes. (A question that can rise from this is that of the possibility of comparing two, or more, valid states. Or, if valid can be put on a scale of *more* or *less*. I will talk about this question more in 5.3.)

---

<sup>43</sup> This, then, being a "sub-reason" or implication of the main reason which is to situate the individuals so that the, by Rawls, preferred result can be reached.

The danger with arguing like this is that there is an implication here where the peoples distorted judgements are so off, so skewed, that the agreement reached will, indeed, be *more* valid than if using The Veil. This if the judgements of the parties were so distorted so that their self-interest were not to be recognised as such, but something completely different. Still, when weighing in the, aforementioned, self-interested disposition, this outcome seems highly unlikely. It is a strange notion, a logical error, to imagine a person who cares about himself, and only about himself, while at the same time caring just as much about everyone else and one would think the the assumption of rationality will prevent this from happening.

So, not here saying that the people in the bargaining situation *will* try to get round the rule of The Proviso but pointing to the possible ability to do so leads to an answer of the question above: It does not matter if they do try to get around The Proviso or if they do not. If there is an agreement at all it would, with almost certainty, be a less valid basis (than if The Veil had been kept) for a theory of justice, if valid at all!

Keeping The Veil seems the better option here.

### **5.3. Criteria of Impartiality**

This issue ties together with the previous one. As stated before Rawls has an idea that theories of justice, in particular, should live up to the fundamental conditions of impartiality, and The Veil can be seen as a strong criteria for impartiality as it gives no room for biased behaviour on any part while, as we have argued before, The Proviso *does* give room for biased behaviour.

To be clear: Gauthier doesn't argue that The Proviso ensures impartiality. He argues, as we have seen, that it is The Proviso together with the two principals of Minimax and Maximin that will capture *fairness* and serve as basis for justice. For Gauthier, however, fairness does not equal impartiality. His thought is that a valid theory of justice may follow from biased (but fairly so) choices. He does, however, argue that the theory would be approved by an impartial, or ideal, observer. This leaves us with two theories, and two initial bargains, claiming to be both impartial and valid. So in choosing between The Veil and The Proviso as our starting point for a social contract theory we need to try and answer questions about more or less valid, more or less impartial, if we are not satisfied with leaving this question as status quo.

The question of more or less valid seems the hardest. However, where it in the last section (5.2.) rested on some thought on the equality between the individuals from the bargaining position after the agreement, and that thought being: The more equal, the more fair, the more valid. (With an implicit idea that the individuals had, in agreeing, satisfied their

self-interests.) Perhaps it in this section could rest on some intuitive idea about more or less impartial, or more to the point: More or less assured impartiality.

It could, however, be the case that it is not possible to talk about more or less impartial. Nonetheless, it seems fairly unproblematic to talk about strong and weak criteria of impartiality. Few would argue against The Veil being as strong an criteria for impartiality as you are going to find. So, at best The Proviso, or the agreement produced through The Proviso and then approved by the ideal observer (from now on, in 5.3., only referred to as The Proviso), will be able to match it. Still, to do so the burden of proof rests on it.

To be fair, to Gauthier, his whole theory is about laying forward this "proof" but no-matter how good an answer and how heavy the proof, questions like: How can a biased judgement be fair? And, how can an agreement between people bargaining with their own stakes be seen as satisfying any strong (or indeed equalling the strongest) criteria of impartiality? Will appear and soil the impartiality claim.

So, in this case as well, but here using impartiality as a base for more or less valid, it seems less of a problem and more of a "clean" starting point to keep The Veil.

#### **5.4. Does Fairness Imply Strong Impartiality?**

Rawls wants to ensure fairness.<sup>44</sup> He wants the individuals in The Original Position to be fairly situated. He wants the agreement reached in the bargain to be fair. And beyond this, he also wants the agreement to lead to a fair society. All this is to be achieved with the aid of The Veil, a strong (if not the strongest) criteria of impartiality.

Now, in the last section (5.3.) I "choose" The Veil (over The Proviso) on a basis of strong versus weak criteria of impartiality. This could, however, be a mistake. It is also Gauthier's last resort: To say that Rawls require too much; we do not need impartiality (strong nor weak) to ensure fairness.<sup>45</sup>

It is really beyond the scope of this essay to examine if this is the case. However, a glance at the "conflict" is still needed.

Now, it seems clear that fairness does not imply impartiality of any kind. Nor does an impartial choice (for example) ensure a fair outcome. So, *does* Rawls require too much?

---

<sup>44</sup> Here setting aside what Rawls actually mean when he writes about it, in the context of *Justice as Fairness*, (a discussion that in itself would take size of an essay equivalent to this one, at least) and settling for a more common, or intuitive, understanding of the concept fair. This, since it to me seems clear that the Rawlsian meaning stems from this concept.

<sup>45</sup> This is not to be confused with the idea that Gauthier thinks his theory, and the decision making there in, will be deemed valid by an impartial, or ideal, observer. The idea here is that we do not need impartiality as a concept in it self to reach fairness.

When Rawls talks of impartiality he talks of it as a result of The Veil but, not only The Veil, there is also the concept of rationality which lead to self interested (mutually disinterested) individuals. Now, *this* kind of impartiality seems to imply fairness. The choice you make is to serve yourself but, not knowing who you are you serve all (and so does everybody else). It is hard to see how this would not be deemed fair once The Veil is lifted.

Gauthier, on the other hand (and as we have stated) talks of impartiality more as a consequence of his theory. But, still claims to capture fairness. He arrives at this fairness through a fairly dense theoretical jungle. Taking the long road of explaining every step of the fully informed individuals instead of just slapping ignorance on them.

Still, somewhere in Gauthier's retreat, in claiming that Rawls require too much, there seem to be a thought towards the acceptance of degrees of fairness. This because one, I doubt Gauthier would say that the outcome in Rawls' theory is unfair. Two, Gauthier thinks the outcome in his theory is fair. And three, the two outcomes are very different (for example regarding the distribution of utility).

This leaves us with two kinds of fairness. One which is ensured through impartiality (and self interest) and, on my account, is very hard to question. And another where every step of the way towards fairness can be questioned (but still here saying that fairness is reached).

While fairness may not imply impartiality it now, to me, seems the better alternative to introduce it as a lead up to what is fair rather than the other way around. This because there seems to be a huge risk in leaving impartiality out. So, in my use of impartiality as a way to distinguish between The Veil and The Proviso, maybe I'm not mistaken after all.

## **6. Summary**

This essay started with the, in some manner, empty question: Is fully informed better than cast in oblivion? The reason for this question was wanting know what would happen if Rawls had not imposed his knowledge restrictions on to the individuals in the initial bargaining situation, his Original Position.

An account for a social contract theory that allows the individuals in the initial bargaining situation to be fully informed was provided by David Gauthier through his book *Morals by Agreement*.

The empty question could now be filled with substance and the main question to answer came to be: Does Gauthier's account of his Initial Bargaining position prove a better start to a social contract theory than Rawls' account of his Original Position, and in that cater to the intuitive idea that fully informed is better than cast in oblivion?

The examination was begun by presenting Rawls' Veil, giving his reasons, showing which restrictions on knowledge that it imposes and presenting arguments for it. Followed, then, by letting Gauthier take the stage: presenting his theory through his five concepts of The Morally Free Zone, the principles of Minimax and Maximin, Constrained Maximization, The Proviso, and the Archimedean point.

At the end of the segment the conclusion reached, and argued for, was that The Proviso was the main point of difference between the two theories (apart from the difference of Gauthier's principals of Minimax and Maximin and Rawls' two principals of justice, a difference not relevant for the purpose essay) and that The Proviso had its adversary in Rawls' The Veil.

A further examination of The Proviso showed that it was, indeed, a moral norm that influenced which agreements could be struck in the Initial Bargaining Position and in that undermining Gauthier's project of founding a theory of morality on mere rationality.

After this conclusion the following segment was used to, in a speculative manner, see if the theory could be used anyway, now seen as a coherentist theory and with The Proviso constituting a rule throughout the bargaining process.

The Veil was compared with the new reading of The Proviso (and Gauthier's three first three concepts as presented in 3.1.) from two angles; choice from distorted judgements and criteria of impartiality. The idea here was to pick out The Veil from two of it's purposes and then replace it with the new reading of The Proviso and see if it does the job as good as The Veil, and in that parting way for full information and shutting the door on knowledge restrictions. Conclusions from both angles are, for the new reading of The Proviso, negative, as it seems The Veil does a better, or in the case of impartiality, perhaps equal, job.

## **6.1 Conclusion**

The final conclusion of this essay comes in form of an answer to this essay's main question: No. David Gauthier's account of his Initial Bargaining position does not prove a better start to a social contract theory than Rawls' account of his Original Position. And my intuitive idea that fully informed is better the cast in oblivion does seem wrong in the light of the conclusion of this essay and in regards to Rawls versus Gauthier.

I will, however, not be distraught. I'll take my intuition and I'll find a challenge for Rawls elsewhere.

## Bibliography

- Gauthier, David. 1986, *Morals by Agreement*, Oxford: Clarendon Press
- Rawls, John. 1971, *A Theory of Justice*, Cambridge, MA: Harvard University Press

## Other Resources

- Ashford, Elizabeth and Tim Mulgan, "Contractualism", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2008/entries/contractualism/>>.
- Cudd, Ann, "Contractarianism", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2008/entries/contractarianism/>>.
- D'Agostino, Fred and Gerald Gaus, "Contemporary Approaches to the Social Contract", *The Stanford Encyclopedia of Philosophy (Winter 2008 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2008/entries/contractarianism-contemporary/>>.
- Freeman, Samuel, "Original Position", *The Stanford Encyclopedia of Philosophy (Spring 2009 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2009/entries/original-position/>>.
- Heath, Joseph, *Envy and Efficiency*, 2009.
- Jollimore, Troy, "Impartiality", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2008/entries/impartiality/>>.
- Wikipedia contributors, 'Archimedean point', *Wikipedia, The Free Encyclopedia*, 9 April 2009, 23:02 UTC, <[http://en.wikipedia.org/w/index.php?title=Archimedean\\_point&oldid=282867375](http://en.wikipedia.org/w/index.php?title=Archimedean_point&oldid=282867375)> [accessed 25 May 2009]