

PIERPAOLO MARRONE

*United Acts of Happiness*

ABSTRACT:

In this paper, I deal with David Gauthier's interpretation of Kant's practical reason. Gauthier argues that happiness plays a unifying role in practical reason and that this function must be seen as a clear hint of the unity of reason. I discuss some suggestions of Gauthier and argue that they are not plausible as interpretation of Kant's practical philosophy.

KEYWORDS:

I. Kant, D. Gauthier, practical reason, rationality, happiness

PIERPAOLO MARRONE \*

## *United Acts of Happiness*

1. *Introduction* – 2. *Happiness and choices* – 3. *Unity* – 4. *Strategic happiness*.

### 1. *Introduction*

David Gauthier has often interpreted his own intellectual work as a response to the challenges of the philosophical tradition. The figure of this Canadian philosopher is interesting in many respects. From a theoretical point of view, his work on the social contract is unanimously considered of the greatest importance<sup>1</sup>. Gauthier's use of the concept of *constrained maximization* is considered as a brilliant response to the problem of cooperation; his generalization of the prisoner's dilemma is held as highly original and thirty years after his formulation still arouses intense discussions<sup>2</sup>. From the point of view of the philosophical style Gauthier is particularly interesting because he has associated his name with some original interpretations of eminent philosophers of

\* Professore Associato, Università di Trieste. E-mail: marrone@units.it.

<sup>1</sup> GAUTHIER 1986.

<sup>2</sup> See *Ethics*, Vol. 123 Numero 4, 2013, with a Symposium on *Morals by Agreement*.

the past: Hobbes<sup>3</sup>, Hume<sup>4</sup>, Rousseau<sup>5</sup>, Kant<sup>6</sup>. His interpretations are not philological and Gauthier does not even seem very interested in surveying the secondary literature, not even the most important or most recent. As he says in the recent book on Rousseau, there is no need to justify what is a personal reading. However, I think we have to be suspicious of this autobiographical and minimalistic consideration. The autobiography is often the road to self-absolution, even when it takes the shape of a reassuring minimalism, as in this account of Gauthier's motivation that led him to his personal reading of Rousseau. The opportunities for reading the philosophers of the past, that Gauthier questions, are rooted in a more general philosophical strategy, that is to seek confirmation of the theories that he defends or to test them in the light of retrospective alternatives. This is the meaning of his confrontation with Hobbes and Hume, in the case of the social contract, or with Kant, regarding his «theory of morality as part of the rational choice theory», according to an expression that had been used by Rawls to describe his theory of justice, but that can correctly describe Gauthier's approach to morality and political cooperation, too.

In these pages I will deal with a case that has to do with this last aspect of the philosophy of Gauthier, mediated by its interpretation of Kant, as it has been issued in his essay *The Unity of Reason: A Subversive Interpretation of Kant*<sup>7</sup>.

<sup>3</sup> GAUTHIER 1969.

<sup>4</sup> GAUTHIER 1990.

<sup>5</sup> GAUTHIER 2006.

<sup>6</sup> GAUTHIER 1990.

<sup>7</sup> GAUTHIER 1990, 110-126.

## 2. *Happiness and choices*

The intention of Gauthier is interpretative in a sense admittedly problematic, and certainly not for the first time. For example, in the case of Hume, Gauthier intended to show that Hume is not akin to the utilitarian tradition, but he should be read as a contractarian thinker. In the case of Kant, the philosophical intention is certainly not less original, since the intention is to give a new reading of Theorems I and II of the *Critique of Practical Reason*, which is openly subversive. Gauthier, in fact, say he does not pretend to give an interpretation that corresponds to the current standards of adherence to the text: rather than overthrowing an interpretative tradition, he wants to subvert even what Kant himself thought, outlining, instead, what Kant would have had to think. «A reinterpretation exploits subversive elements and present ideas in a text to lead the reader in a direction manifestly different from and even opposed to that in which the author seeks to go. But this new direction is not randomly selected.» The interpretative ambition does not stop here: Gauthier believes that this strategy should lead to a better understanding of the rationality inside moral behavior. «Fidelity to the text is sacrificed, not for the mere commentator's delectation, but in a philosophic causes».

Gauthier assumes the philosophical background of practical philosophy of Kant, in order to reject the two theorems of practical reason, in the light of some prerequisites of kantianism which are part of the *Critique of Pure Reason*. It is clear that we are already beyond Kant himself, since this is a interpretative hypothesis that should be rejected in the light of the first lines of the *Critique of Practical Reason*.

«Why this *Critique* is not entitled a *Critique of Pure Practical Reason* but simply a *Critique of Practical Reason* generally, although its parallel with the speculative seems

to require the first, is sufficiently explained in this treatise. I has merely to show *that there is pure practical reason*, and for this purpose it criticizes reason's entire *practical faculty*. If it succeeds in this it has no need to criticize the *pure faculty itself* in order to see whether reason is merely making a claim in which it presumptuously *oversteps* itself (as does happen with speculative reason)<sup>8</sup>.

The problem, however, is not taking seriously this parallelism, which, if read literally, produces similarities and not intersections, but «to lead you to a truer understanding of the role of reason in ethics. Kant is to serve as an unwilling recruit in this task»<sup>9</sup>.

The two theorems of practical reason are intended to eliminate any empirical determination as a determining motivation of moral action. The first theorem, in fact, says, «[a]ll practical principles that presuppose an *object* (matter) of the faculty of desire as the determining ground of the will are, without exception, empirical and can furnish no practical laws»<sup>10</sup>. The second theorem states that «[a]ll material practical principles as such are, without exception, of one and the same kind and come under the general principle of self-love or one's own happiness»<sup>11</sup>.

For Kant, from the point of view of morality, it is not possible to unify the manifold material given in the experience with the empirical concept taken from this same experience. Let us think about desire. The desire needs an imaginative representation which builds in our mind the image of a singular object. «For the determining ground of

<sup>8</sup> KANT 1997, 3.

<sup>9</sup> GAUTHIER 1990, 110.

<sup>10</sup> KANT 1997, 19.

<sup>11</sup> KANT 1997, 19.

choice is then the representation of an object and that relation of the representation to the subject by which the faculty of desire I determined to realize the object»<sup>12</sup>. The principle of desire is, in other words, always empirical. It is also difficult to consider it as a principle in the sense of being a rule of unification of the empirical, since we desire a particular object. If unification happens, this happens because in the practical field one must go beyond the empirical use.

«As finite beings, lacking self-sufficiency, we have desires; as rational beings, we unite our conception of these desires in thought and so conceive of happiness as the proper object of desire. Rationality, applied to finitude, necessitates the idea of the satisfaction of all our desires and so necessitates the idea of happiness»<sup>13</sup>.

Gauthier's idea is that, if there is a rational faculty which is the origin of the action, namely the will, and since this is defined by Kant as nothing more than practical reason, we should expect even here some concepts that unify data experience, as in the case of pure theoretical reason. It is surprising, however, that Kant does not provide any of these concepts<sup>14</sup>. Actually, Gauthier thinks that Kant had effectively provide a practical pure concept, but has not had the courage to think it consequently as the unifying principle of moral experience. This concept is happiness and Gauthier asks for a deduction – that is a justification, in kantian jargon –, stemming from the multiplicity of desires. The analogy with the theoretical knowledge is revealing:

<sup>12</sup> KANT 1997, 19.

<sup>13</sup> GAUTHIER 1990, 112.

<sup>14</sup> GAUTHIER 1990, 115.

«As space and time are the conditions of our speculative apprehension of objects, so need, we may suppose, is the condition of our practical apprehension. Apart from need, objects would not be grasped by us as of interest or concern; knowledge of them would be possible but would not dispose us to act in relation to them»<sup>15</sup>.

Gauthier believes that this is a feasible task and that does not give raise to the objections which are usually raised against it. For example, this: we know the desires of the agents from their choices, and these choices reveal different orders of preferences, which should make us doubt about the possibility of unifying them through a practical law.

Gauthier's response to this objection is twofold. (1) It is true that an indefinite number of choices not always reveals a well-defined order of individual preferences, and not only because we cannot access to the intentions of the agent, either directly (introducing us in his head), or indirectly (by deduction from his actions), but also because in the time-sequence of the choices, the strategic coherence of identifying an order of preferences is a requirement that sometimes remains unknown to the same agent. Gauthier thinks that this does not in any way affect the idea of happiness as an unifying concept.

«The actor's desires must be so related that they determine a preferential ordering of the set of alternative possible actions, from which she may then select a maximal element. The familiar ideas of the theory of rational choice correspond to the pure concepts of the will»<sup>16</sup>.

<sup>15</sup> GAUTHIER 1990, 112.

<sup>16</sup> GAUTHIER 1990, 116.

(2) Happiness is not simply given as a natural need for all sentient and rational beings, but as what unifies all desires, that is happiness is the sufficient and necessary reason of the existence of the same concept of desire.

«Happiness, the satisfaction of all desires, is given as the end of action not, as Kant seems to have supposed, by natural necessity, but as the result of the activity of the will, or practical reason, in unifying the manifold of desire to determine a single action. The action so determined takes happiness as its end»<sup>17</sup>.

These quotations can be read as a commentary on what Kant says in the *Groundwork of the Metaphysics of Morals* about the concept of practical law. Kant writes that

«We shall thus have to investigate entirely a priori the possibility of a *categorical* imperative, since we do not here have the advantage of its reality being given in experience, so that the possibility would be necessary not to establish it but merely to explain it. In the meantime, however, we can see this much: that [...] by contrast the categorical imperative alone has the tenor of a practical law; all the others can indeed be called *principles* of the will but not laws, since what it is necessary to do merely for achieving a discretionary purpose can be regarded as in itself contingent and we can always be released from the precept if we give up the purpose; on the contrary, the unconditional command leaves the will no discretion with respect to the opposite, so that it alone brings with it that necessity which we require of a law»<sup>18</sup>.

<sup>17</sup> GAUTHIER 1990, 116.

<sup>18</sup> KANT 2012, 50.

The existential condition of each of us is to have always a variety of needs and desires. These existential objectives do not clearly manifest all together at the same time, but they usually overlap in different circumstances of the life of the agents. Some of these cannot occur together, because they are mutually incompatible, in the sense that the presence of one excludes that of another one. Therefore, it is not possible to identify each of these individual goals with happiness. The identification would be an undue and inordinate simplification. Moreover, we cannot even think that happiness is a desire or a need, among others, since happiness is a goal and a purpose, namely a mode of selection of desires and needs, each of them chosen according to their ability to make us reach that goal. The function of happiness is not for Gauthier simply instrumental, but rather is an ordering function. Needs and the desires lack of this ordering function.

«In a single choice the existence of a maximal element is an analytic necessity; the chosen element may always be interpreted as maximal. But a series of choices may not reveal a single preferential ordering of alternative possible actions. Taken together, they may not express the unified desires of an individual rational actor»<sup>19</sup>.

Happiness is, therefore, that pure concept one should look for in the field of practical reason. This is what Kant have deliberately failed to recognize. In *Anthropology from the Pragmatic Point of View*, however, Kant hints at this ordering function of universal happiness, as linked to the ideal of progress of mankind, when he suggests that if wars testify to a destructive impulse, this does not prevent

<sup>19</sup> GAUTHIER 1990, 116.

rational creatures, with the gradual increase of their culture despite the wars, to offer to mankind for centuries to come the clear prospect of a state of happiness irreversible<sup>20</sup>.

This universal perspective of the human race, however, is certainly not interesting in this circumstance for Gauthier. This kind of unification, by mean of a future regulatory perspective, is postponed sine die, while for Gauthier what is important to happiness is the ability to make sense of the actions of each individual. Indeed, behind this idea of Gauthier there is a whole conception of the individual as a strategic agent. For this strategic condition it is not possible to be an individual who simply satisfies his/her immediate needs and desires, without any other additional unifying perspective (deciding, if possible, to satisfy his/her own current needs and desires is itself ultimately a unifying perspective). Of course there is the problem of knowing what is required for being something that gives unity to a life. It is not sufficient that there is always a goal, because, after all, also sponges pursue their goals with unflinching constancy. Two further conditions, lacking in animals<sup>21</sup>, must be satisfied: (1) the subject must be able to make a semantic representation of this goal, relatively constant over time, and (2) this representation refers to an actor who is also relatively constant in time<sup>22</sup>. These conditions are background necessary conditions to the concept of rational individual, even if they are not sufficient conditions. In this sense, they apply to a concept of individual agent that is a transcendental concept. I think it is clear the analogy with the transcendental unity of apperception. As the concept of 'I think' is necessary to give meaning to our experience –

<sup>20</sup> KANT 2006, par. 88.

<sup>21</sup> See also LEAHY 1994; *contra* GRIFFIN 1994.

<sup>22</sup> GAUTHIER 1988.

that is, to think about the experience as ours, because otherwise we would have, Kant says, representations without thinking –, so it is necessary a unified concept of agent to make sense of our actions. This concept, however, must show something that must be common in every action. This cannot be but an intention. The unity of reason mentioned in the title of Gauthier's essay it is to be understood in this way.

### 3. *Unity*

But this kind of transcendental unity of apperception in action (the unified semantic representation of an individual who is acting) is really a necessary and sufficient condition? Happiness is really the unifying function of desires, and this function must be thought in strategic terms? For Kant happiness was just an accessory condition of any good action. Kant would certainly not have signed the statement of Lev Tolstoj at the beginning of *Anna Karenina*, «Happy families are all alike; every unhappy family is unhappy in its own way», because happiness does not have the unifying power of rational action. As Kant writes:

«Now in a being that has reason and a will, if the proper end of nature were its preservation, its welfare, in a word its happiness, then nature would have hit upon a very bad arrangement in selecting the reason of the creature to carry out this purpose. For all the actions that the creature has to perform for this purpose, and the whole rule of its conduct, would be marked out for it far more accurately by instinct, and that end would have thereby been attained much more surely than it ever can be by reason; and if reason should have been given, over and above, to this favored creature, it must have served it only to contemplate the fortunate

constitution of its nature, to admire this, to delight in it, and to be grateful for it to the beneficent cause, but not to submit its faculty of desire to that weak and deceptive guidance and meddle with nature's purpose. In a word, nature would have taken care that reason should not break forth into practical use and have the presumption, with its weak insight, to think out for itself a plan for happiness and for the means of attaining it. Nature would have taken upon itself the choice not only of ends but also of means and, with wise foresight, would have entrusted them both simply to instinct»<sup>23</sup>.

It seems clear that the unifying power of the reason is given by a personal rational intention and not by the fact that instinct achieves more precisely its target. But, according to Kant, it would not be even a necessary condition. Just think an agent acting in accordance with the principles of act-utilitarianism.

Kant and Gauthier both share the idea of a moral law as a semantic structure that does make sense of the manifold elements of the empirical experience, but for the former this law is something that can exist outside human cooperation; for the latter is something that show the irrationality of maximizing without any restriction. For Gauthier:

«If an actor brings her need-based desires into a single framework, so that they direct her to her maximum overall satisfaction, then she is not inevitably determined to respond directly to each of her needs. As we have seen, the necessity of happiness as the object of desire is not natural but practical, determined by the unifying activity of practical reason. Thus such an actor is autonomous, acting

<sup>23</sup> KANT 2012, 34.

on the basis of the practical law given by her exercise of practical reason. If there is a problem in reconciling practical necessity with natural necessity, or human freedom with causality, then it arises as much for the actor who acts autonomously to maximize her own happiness as for the actor who acts to fulfill the moral law. Kant's emphasis on and concern with autonomy is retained by our reinterpretation»<sup>24</sup>.

But the content of this semantic structure is very different and it is thanks to a kind of *petition principii* that Gauthier believes that he has found the theoretical exit to the kantian skepticism about the ability to find a moral law in the material principles of the faculty of desire: the law of practical reason is not so much in the maximization of happiness, as in the ability to pose constraints on the pursuit of happiness. That constraint that, if anything, for Kant would have been only an additional ancillary condition to the fulfillment of duty (acting morally, the agent becomes worthy of happiness), for Gauthier becomes a necessary condition to release the rational actor from the contingency of satisfaction of desires.

Kant would have considered this move neither persuasive, nor conform to the universality that we have to ask to the moral law. He thought that a law of practical reason based on happiness would produce practical destructive effects. If everyone were entitled to pursue his own well-being, there could be no agreement between the agents, if not in an entirely contingent way. According to Kant, even when an agreement is produced, this would not be necessarily a desirable outcome. The example given by Kant is that of a married couple who decide to break the

<sup>24</sup> GAUTHIER 1990, 117.

union. The agreement of the two will seems here perfect, but the result will not certainly produce harmony.

For Gauthier, as for Kant, neither utilitarianism nor selfishness can provide a solid basis for moral behavior. This solid base can instead be obtained if the agents are able to provide a rational rule on the conditions of choice and a rational rule on the outcomes. Of course, Gauthier believes to be in possession of these two rules which are as follows: (1) everyone should maximize their own happiness, given the actions of other agents, and (2) everyone should maximize their own happiness, given the amount of happiness received from other agents<sup>25</sup>. These rules are not only extremely demanding, but they are inapplicable, because they can sketch their effective content just taking the point of view of an omniscient observer. In fact, there is something extremely indeterminate in their formulation, and it seems very difficult to specify some conditions in such a way as to make them manageable in the moral calculus. I especially refer to the notion of 'other'. Even when with 'other' we must understand 'other human beings', the kind of being we share the ability to form semantic representation with (while this ability is not present in animals), how to use this notion remains completely undetermined, because it remains entirely undetermined how it must be further specified. 'Other' cannot be regarded as a primitive term in any moral conception, even in those extremely formalized. Who is the other? On the one hand the response appears to be extremely simple: the other is a human being who does not occupy my same physical space and that is capable of cognitive performance comparable with mine, but, in another sense, this is only the condition of an effective communication and not the sufficient condition of any moral cooperation.

<sup>25</sup> GAUTHIER 1990, 119.

Kant could agree that here we have some unifying rules of desires, but probably he would object that it is not clear whether these rule specify normative or descriptive conditions. Are they really rules that everyone must follow, or rather are they rules that everyone already follows? Putting together the words 'action', 'rational', 'rules' does not produce something morally relevant from the point of view of Kant, because these words could simply identify general laws that are maxims of action, but not moral a priori laws. But Gauthier's rules may tell us something else, that is the prevalence of the order of individual preferences, and the very fact that this order is to be understood as a belief about the future course of his/her actions and the future course of actions of the actors, which presumably are affected by the expected results of my actions. Kant would say, however, that they are nothing more than hypothetical imperatives.

#### 4. *Strategic happiness*

These considerations would seem to substantiate the kantian idea that what form our idea of happiness is taken from the senses, from the material given to us by our experience of the world: and this, according to Kant, means to deny that it can dictate any valid law for practical reason, since a practical law must be given to the practical reason only by itself. Therefore, when Kant write that

«Thus, it is found, for example, that we can find satisfaction in the mere exercise of our powers, in consciousness of our strength of soul in overcoming obstacles opposed to our plans, in cultivating our talents of spirit, and so forth, and we correctly call these joys and delights more refined because they are more under our control than others, do not

wear out but rather strengthen feeling for further enjoyment of them»<sup>26</sup>,

he means that we are playing with empirical feeling to a higher level. So again, it is not possible to derive practical moral laws from such mental states.

For Gauthier, the error of Kant does not consist in adopting too narrow criteria for the formulation of laws in the field of action, but in conceiving happiness simply like other desires<sup>27</sup>. Happiness, on the contrary, is neither a desire among others nor the particular object of a desire, but the general strategy in the actions of the subject. We want to achieve some purpose, possess some object, succeed in a certain task because these tasks give us happiness, or satisfaction or utility. The use of three different terms does not seem to change the substance of the problem: they show in each case that the unification is that of a desired purpose through a multiplicity of forms.

Gauthier's thesis is that happiness is the unifying element of desires, so by its very structure when we look at happiness we are looking at something other than particular wishes, desires, goals. I think that there is always some circularity at stake in the arguments that are concerned with happiness, desires, needs. This should not be surprising that much since the concept of happiness is tightly woven with those of need and desire. We want something and/or we need to reach a certain goal because it gives us happiness, but the happiness that we reach – if and when we reach it – is the fulfillment of initial need and desire, and forms an important part of our impulse to act. It could be argued that this circularity is unavoidable and inescapable, because it is

<sup>26</sup> KANT 1997, 21.

<sup>27</sup> GAUTHIER 1990, 115.

part of our way of conceptualizing needs, desires, happiness. However, one should not think that we are talking about a single object. The idea that happiness is a unifying strategy might suggest that, after all, happiness is a single object. But from a strategic point of view, happiness looks like more as a process than as an object. From his point of view, however, Kant had strong reasons to refuse an unification drawn from the empirical world to shape practical moral laws. As a matter of fact, the skeptical objection managed by Kant, obviously for not skeptical purposes, can be repeated against Gauthier: happiness is an empty concept that can be filled in any way. Kant would have had very strong doubts that taking a strategic point of view – true pillar of the criticism of Gauthier and his idea of unity of reason – can give rise to a single rational outcome and not to a lacking of moral normativity, masked by the supposed uniqueness of a procedure in action.

But perhaps at this point we must ask what are the underlying reasons that justify Gauthier's subversive reinterpretation of Kant, which overlaps theoretical reason with practical reason. One motivation is clearly expressed in another paper, that is Gauthier's *Why Contractarianism?*<sup>28</sup>. According to Gauthier, moral theories, that want to answer the skeptical challenge, can be divided into two groups. On the one hand, there is the attempt to justify morality from the inside. For this view, morality should provide reasons for action, which are independent from desires, personal goals, interests. But what other reasons has a person to act?

«From the standpoint of the agent, moral considerations present themselves as constraining his choices and actions, in ways independent of his desires, aims, and interests. [...]

<sup>28</sup> GAUTHIER 1991.

For it reveals clearly what is in question – the ground of constraint. This ground seems absent from our present world view. And so we ask, what reason can a person have for recognizing and accepting a constraint that is independent of his desires and interests? He may agree that such a constraint would be morally justified; he would have a reason for accepting it if he had a reason for accepting morality. But what justifies paying attention to morality, rather than dismissing it as an appendage of outworn beliefs? We ask, and seem to find no answer»<sup>29</sup>.

The kantian practical reason is the most obvious example of this way of founding morals. John Rawls's theory of justice is another kantian way to defend a kind of normativism linked to principles and to a specific moral psychology. Gauthier thinks that internalism faces a foundational crisis, but that that crisis is not necessarily the crisis of morality. What we need is for Gauthier an external foundation, that is a theory of constraints. This means that if there is a practical reason, it must be possible, in principle, acting in such a way that the conditions of cooperative rationality are fulfilled. This means that, again, in order to fulfill this task, you must define the practical rationality independently from morality<sup>30</sup>.

The kantian interpretation of Gauthier is clearly subversive, but for reasons different from those put forward by the author. More simply, it is a deliberate misunderstanding of the practical philosophy of Kant. For the purpose of Gauthier is to show that there is not discontinuity between theoretical reason and practical

<sup>29</sup> GAUTHIER 1991, 16.

<sup>30</sup> Even if GAUTHIER 2013 modifies the initial setup of *Morals by Agreement*.

reason, and that the line of continuity must be sought both in happiness and in the principle of minimax relative concession, according to which

«each bargainer will be most concerned with the concessions that she makes from her ideal outcome relative to the concessions that others make. If she sees her concessions as reasonable relative to the others, considering that she wants to ensure as much for herself as she can while securing agreement (and thereby avoiding the zero-point: no share of the cooperative surplus) and subsequent compliance from the others, then she will agree to it. What would then be the reasonable outcome? Gauthier argues that it is the outcome that minimizes the maximum relative concessions of each party to the bargain»<sup>31</sup>.

How this strategy can be set apart from rule-utilitarian is not easy to understand, especially in light of the notations on Kant in *Morals by Agreement*. For example, when Gauthier discusses the choice made from the Archimedean point of view – which is his alternative thesis to the original position of Rawls – he also takes the opportunity to exhibit his distance from the kantian position of Rawls himself. For Gauthier although the ignorance of their identity precludes any manifestation of positive bias in agents' choice, this is insufficient to ensure an equal rationality. The impartial selection is not the absence of concern for those who are involved, but the presence of an equal concern. And this equal concern is ensured by the purpose of the actor's maximizing ideal. Although he cannot identify with anyone, everyone can identify with him. The impartiality of his choice is then showed in the character fully representative of

<sup>31</sup> CUDD 2012.

his choice<sup>32</sup>. But these considerations lead us not in the neighborhood of Kant, but rather in humane suburbs, since they are fully consistent with Hume's idea that reason is and must necessarily be a slave of the passions. As a matter of fact, Gauthier rejects, in his major work, the contiguity with the kantian position, since for Gauthier the kantian indifference to the passions is completely foreign to his subject<sup>33</sup>. Is there a discrepancy between the article from which we started, his idea to subvert Kant through the use of happiness as a strategic tool of reason, and the major work? The answer must be negative. According to *Moral by Agreement*, agents must act to maximize their own constrained utility. Since this is the rationale of their action, the discussion of the unifying function assigned to happiness assumes the entire structure of Gauthier's contractual theory, a structure that he cannot find in Kant. For Gauthier, there are three levels in contractual theory: (a) the claim that rational and moral constraints are internal constraints. These constraints, once defined in an impartial way, must be included as part of the theory of rational choice, (b) the introduction of a special procedure, the purpose of which is to show that it is rational to constraint when others are doing the same, and (c) the assumption of the principles that govern the structure of impartial constraints: the minimax relative concession, and the lockean condition.

Of these three conditions, Gauthier can find only the first in Kant. The others are at odds with the idea of practical kantian reason. In particular, if we assume the idea of a kantian moral duty, the conditions (b) and (c) would be like playing consequentialism in a deontologic frame.

<sup>32</sup> GAUTHIER 1986, 236.

<sup>33</sup> GAUTHIER 1986, 236.

This incompatibility between Gauthier and Kant is not due to what Kant would have to think and has not thought, but rather to the function of morality, which in Kant is not linked to the passions, but that works as a function of unification of the manifold, which is represented by the infinite number of opportunities to exercise of morality. However, this manifold is always given also as unity, that is as duty. So it follows that Kant cannot properly speak of moral error, but only of error in the action – I may be wrong in the action believing that my result will be consistent with my moral intention, but I cannot go wrong in my intention, if the intention is moral –. In Gauthier, however, the internal constraints of morality are the mirror of the external ones and are justified on their basis, since they are not really independent. They are subject to a process of negotiation and calculation. In the world of social constrained cooperation they are certainly inevitable, but into the inner world of duty built by Kant one cannot find space for them. For this reason, for Kant happiness remains an empirical element located outside the horizon of moral duty. Kant does not run into a contradiction thinking that acting morally makes you worthy of happiness. He thought that for deserving this it is necessary postulating both the immortality of the soul and the existence of God, for moral action requires an infinite time to be fully consistent to the formal structure of moral duty and for approaching asymptotically to the union of virtue with happiness, represented by the idea of God.

Gauthier's unity of reason must conceal all these elements, since they cannot find space in a rational choice theory. So the conclusion that is perhaps more prudent to draw is that Gauthier's reading of Kant is, therefore, more an intelligent and intentional misunderstanding than a subversive interpretation. A misunderstanding intending to show Gauthier's own theory of the unity of reason and not the strength of the kantian division between theoretical reason and practical reason.

## References

- CUDD A. 2012. *Contractarianism*, in ZALTA E.N. (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2012 Edition). Available from: <http://plato.stanford.edu/archives/fall2012/entries/contractarianism/> (accessed 10/09/2013).
- GAUTHIER D. 1969. *The Logic of Leviathan*, Oxford, Oxford University Press.
- GAUTHIER D. 1986. *Morals by Agreement*, New York, Oxford University Press.
- GAUTHIER D. 1988. *Morality, Rational Choice, and Semantic Representation: A Reply to my Critics*, in «Social Philosophy and Policy», 5(2), 1988, 173-221.
- GAUTHIER D. 1990. *Moral Dealing*, Ithaca, Cornell University Press.
- GAUTHIER D. 1991. *Why Contractarianism?* in VALLENTYNE P. (ed.), *Contractarianism and Rational Choice*, New York, Cambridge University Press, 1991, 15-30.
- GAUTHIER D. 2006. *Rousseau: the Sentiment of Existence*, Cambridge, Cambridge University Press.
- GAUTHIER D. 2013. *Twenty-Five On*, in «Ethics», 123(4), 2013, 601-624.
- GRIFFIN D. 1994. *Animal Minds*, Chicago, Chicago University Press.
- KANT I. 1997. *Critique of Practical Reason*, translated and edited by Mary Gregor, Cambridge, Cambridge University Press.
- KANT I. 2012. *Groundwork of the Metaphysics of Morals*, translated and edited by Mary Gregor, Cambridge, Cambridge University Press.
- KANT I. 2006. *Anthropology from a Pragmatic Point of View*, translated and edited by Robert B. Loudon, Cambridge, Cambridge University Press.
- LEAHY M. 1994. *Against Liberation*, London, Routledge.