Moral Motivation in Kant

Konstantinos Sargentis, University of Crete

The problem of moral motivation, which is central to Kant’s account of autonomy, could cover a wide range of issues from many different perspectives. Nonetheless, in Kant’s main discussion of moral motivation it takes a specific form, in that it refers to the ‘fact’ that pure reason becomes practical and provides an incentive for moral action. Kant identifies this incentive with respect for the moral law, yet his account of respect is so condensed that it may be interpreted in diametrically opposing ways; in particular, the fact that respect is presented as both affective and intellectual gives rise to two different readings of Kant’s theory of moral motivation, that is, the intellectualist and the affectivist reading. At the same time, it makes it possible to interpret Kant as either an internalist or an externalist. Both questions, that is, the question concerning intellectualism or affectivism and the question concerning internalism or externalism, are currently under debate in contemporary literature. This paper is a contribution to this debate, since, on the one hand, I support the affectivist reading and, on the other hand, contrary to the view of most commentators, I argue that Kant is an externalist. Moreover, by bringing the two questions together, I attempt to show that Kant is an externalist concerning motivation because he is an affectivist.

The present paper is divided in six parts. In the first part, I examine the distinction between internalism and externalism regarding the issue of motivation and I present my own understanding of it. Moreover, I offer a further distinction

1 I would like to thank Wolfgang Ertl, Paul Gilbert, and two anonymous reviewers of the Kant Studies Online for their helpful comments on earlier versions of this paper.
2 For example, see the contributions in Klemme/Kühn/Schönecker 2006.
between the cognitivist and the non-cognitivist versions of internalism and externalism. In the second part, I place the development of Kant’s theory of motivation within the framework of that distinction and I argue that, despite the differences between the various phases of this development, Kant was a cognitivist externalist at least throughout the Critical period. In the third part, I examine two possible objections concerning the scope of Kant’s theory of moral motivation and I argue that those objections evaporate as soon as one focuses on Kant’s main philosophical concern. Furthermore, I identify this scope with the role the feeling of respect for the moral law plays in Kant’s ethics; thus, in the fourth and fifth parts I proceed to the analysis of the nature and the function of respect respectively. Moreover, in the fifth part, in order to illustrate the function of respect as an incentive for moral action, I draw on the controversy between intellectualists and affectivists and I argue that the affectivist reading of Kant’s theory of moral motivation is the most appropriate one. Finally, in the sixth part I reach the conclusion that Kant’s affectivism proves him to be a cognitivist externalist.

In the modern debate the question regarding the practical character of morality has led to a distinction between internalism and externalism. Despite the fact that there exist many versions of that distinction, it is crucial to distinguish between internalist and externalist views concerning reasons, and internalist and externalist views concerning motivation. In the present paper I will deal exclusively with the problem of internalism and externalism concerning motivation. As Thomas Nagel puts it:

Internalism is the view that the presence of a motivation for acting morally is guaranteed by the truth of ethical

---

3 On the different versions of this question, see Brink 1989, 37 ff.
 propositions themselves. [...] Externalism holds, on the other hand, that the necessary motivation is not supplied by ethical principles and judgments themselves, and that an additional psychological sanction is required to motivate our compliance.  

Although Nagel’s distinction here is quite clear, the distinction between internalism and externalism regarding the issue of motivation is often employed by philosophers in an ambiguous way.  

And, notwithstanding the clarity of his distinction, Nagel himself uses it to present Kant as an internalist, which is, according to my view, not supported by his above cited explanation. To be sure, most commentators in the literature consider Kant to be an internalist – that is, besides Nagel, interpreters such as, Falk, Korsgaard, Allison, Guyer, and Dancy.  

I believe that behind this characterisation of Kant there is a certain reluctance to make an unambiguous distinction between reasons for action and motivational reasons. Thus, it is usually the case that when such scholars discuss the question of motivational reasons they are really referring to the problem of reasons for action. David Brink raises this point with reference to Falk and Nagel amongst others:  

Internalism about motives and internalism about reasons are often confused [...] Moral considerations are said to be motivational, even if they do not actually motivate, so long as they have the capacity to motivate. But what these writers usually mean when they say that $x$ has the capacity to motivate is that a rational person would be motivated to do $x$ or that an agent would be motivated to

---

5 Cf. Brink 1989, 40 f.
6 See Nagel 1970, 11 f.; Falk 1947-1948; Korsgaard 1996, 311-334; Allison 1990, 238, 241, 283; Guyer 2000, 136 (yet, as Guyer points out, this was not always Kant’s position, see ibid., n. 8); Dancy 1993, 7 ff.
7 See Brink 1989, 38 ff., 51.
do x if she were rational, and this is just to say that one has reason to do x.\(^8\)

Now, what ensues from this ‘confusion’ is that when those scholars examine Kant’s case, they focus on the fact that the ultimate source of moral motivation is pure reason (rather than the human needs that stem from sensibility), which is how the interpretation of Kant as an internalist is itself motivated.

An exception in the literature is Frankena who, despite the fact that he follows the same route, considers Kant to be an ‘externalist’.\(^9\) Frankena describes externalists as “those who regard motivation as external to obligation” and internalists as “those who regard motivation as internal to obligation”.\(^10\) He further explains that “externalists insist that motivation is not part of the analysis of moral judgments or of the justification of moral claims“, whilst, according to internalists, motivation is “involved in the analysis of moral judgments and so is essential for an action’s being or being shown to be obligatory“.

Frankena illustrates that the opposition between the two parties does not concern the question whether morality is to be practical, since both agree on this point. Rather, as he argues, “the question is whether motivation is somehow to be ‘built into’ judgements of moral obligation”.\(^12\) In my view, the latter explanation is very appropriate, yet when Frankena applies it to Kant’s case he fails to make any reference to the theory of respect. On the contrary, by ascribing to Kant the thesis that “obligation represents a fact or requirement which is external to the agent in the sense of being independent of his desires or needs”,\(^13\) he counts Kant as a paradigmatic externalist.

It is worth mentioning that Christine Korsgaard has examined this very point. Interestingly, she finds Frankena’s
view “a surprising difference of opinion”, since, as she states, “Kant himself took the categorical character of the imperative and autonomy of the moral motive to be necessarily connected”. However, it is my view that Frankena here has no intention of denying this essential connection between the categorical moral command and the ‘autonomous’ moral motive, since he does not even consider it. Rather, in discussing an argument by G. C. Field against externalism, he merely expresses the idea (implied in Field’s criticism) that according to Kant the demands of morality (as such and as motivating factors) do not rest on the sensible nature of man and his needs as a sensible being: they are ‘external’ to them. It is in this sense, as it seems to me, that Frankena ascribes externalism to Kant. Thus, I do not find this as surprising as Korsgaard does. Yet, the real problem with Frankena’s view is that he counts Kant as an externalist specifically for this reason. And this is problematic, because, although what may be drawn directly from the independent character of moral obligation is Kant’s rationalism (which is also implied in Korsgaard’s internalist interpretation of Kant), the externalist requirement is not yet fulfilled. In order to attribute externalism (or internalism) to Kant, one must focus on what Kant himself explicitly says about moral motivation, that is, one must focus mainly on his theory of respect for the moral law as an incentive for moral actions.

Thus, since I have yet not found in the literature a satisfying and ‘fair’ way of placing Kant’s theory of moral motivation in the debate between motivational internalists and externalists, I need to further clarify my own understanding of the problem. In order to do that, I will employ a second distinction, that is, the distinction between the cognitive and the non-cognitive versions of internalism and externalism. To start with, cognitivism is the view that moral judgements express cognitive attitudes or states of mind, such as belief; insofar as they are cognitive, they have a certain conditional truth property depending on whether moral facts exist and on the way moral facts are. Non-cognitivism, by contrast, is

---

the view that moral judgements express non-cognitive or conative attitudes, such as desire, which lack the above-mentioned truth property. Now, if the practicality of moral judgements, that is, their capability of guiding, determining or influencing human action, is to be shown, the non-cognitivist is better off, since desires and the relevant non-cognitive states of mind are by their very nature dynamic factors. Non-cognitivists can very easily endorse internalism, since in their case it is certainly obvious that moral judgements can motivate in virtue of the non-cognitive (conative, dynamic) state they express. On the other hand, the cognitivist has a quite hard task to accomplish, because s/he has to show how a judgement without expressing, or being internally connected with, a dynamic state of mind could provide motivation for action. Therefore, the cognitivists have to find out another link to motivation, and in their attempt to fulfil this task they face two alternatives. As Dancy puts it, “they can be internalists, holding that moral judgements express peculiar beliefs which, unlike normal beliefs, cannot be present without motivating. Or they can be externalists, holding that moral judgements express beliefs which rely on the presence of an independent desire if they are to motivate.” For the sake of clarification, it must be noted that in the last case it is the independent desire itself which supplies the moving force to action and not the belief which relies on its presence and is expressed in the moral judgement.

Based on what was discussed above, I can now turn to my own understanding of the distinction between internalism and externalism with regard to motivation. My departure point as regards the present paper consists in this

15 Hume is perhaps the best known non-cognitivist internalist.
16 Dancy 1993, 3. For a slightly different explanation of the distinction between cognitivism and non-cognitivism, see Smith 2002, 4 ff. According to Smith, each of those two notions is the psychological side of the “objectivity of moral judgment” (cognitivism) and of the “practicality of moral judgment” (non-cognitivism), which are constitutive of the “moral problem”, as he calls it: “Can we reconcile the objectivity and the practicality of moral judgement with the standard picture of human psychology that we get from Hume? This is what I call ‘the moral problem’” (14).
specific understanding, which, in short, takes the following form: (a) If someone holds that moral knowledge (in the form of judgements) expresses cognitive states of mind, such as beliefs, and that in doing so it necessarily motivates action, then s/he is a cognitivist internalist. 17 (b) But, if someone believes that moral knowledge expresses non-cognitive attitudes, such as desires and feelings, and that it motivates exactly because it is the expression of such conative attitudes, s/he is a non-cognitivist internalist. These are the two ways of being an internalist in the theory of moral motivation, that is, of believing that moral knowledge per se motivates. But, on the other hand, if one believes that motivation is exclusively a matter of a non-cognitive (and, therefore, external to knowledge) state, which is not a condition of moral knowledge, then s/he is an externalist: knowledge is one thing and motivation is another. Here again there are two ways of being an externalist. (c) If someone holds that the moral judgement, to which motivation is external, expresses a cognitive state and that in doing so it is relevant to motivation, s/he is a cognitivist externalist. In order to be relevant to motivation in that case, knowledge must contribute to it indirectly. Therefore, there must be a link between the external motivating factor and knowledge, but only provided that the motivating state is not a condition of knowledge (otherwise, this would lead us back to internalism). Such a connection is possible only if the motivating factor is a non-cognitive, conative state which is caused or produced by knowledge. Although in that case the conative state depends on knowledge, motivation itself is independent of it (it is ‘external’ to knowledge). For, what moves is the conative state as a non-cognitive or affective state and not the knowledge by which it is produced. (d) But, if one believes that moral judgements express non-cognitive attitudes which, despite being conative, cannot (because of a certain weakness or insufficiency) move towards the corresponding action, and that in order to do so they need an extra conative

17 The Socratic doctrine that knowledge is sufficient for virtue and that akrasia is not possible could be an example of cognitivist internalism.
psychological factor, then s/he is a non-cognitivist externalist.\footnote{I must admit that it is relatively hard to identify one of the already known moral theories as a non-cognitivist externalist moral theory. Above all, it is hard to see how non-cognitivism does not necessarily lead to internalism, thus leaving open the possibility of externalism. Yet, I think it is possible to imagine the workings of such a theory. More specifically, that could be the case of a non-cognitivist theory, in which moral motives (the conative states expressed in moral judgements) provide some but not sufficient motivation for action, and thus they need the contribution of other, non-moral motives to be effective in sufficiently motivating action. In other words, it could be the case for a theory, in which the possibility of overdetermination of action is needed as part of the theory for a comprehensive explanation of the possibility of moral motivation. Following Brink (1989, 41), one could describe this principled motivational insufficiency of moral judgments as a weak form of internalism. In that case, and since I take internalism to be necessarily of a strong, and not of a weak, form, since in my view internalism holds that moral judgements provide sufficient motivation, non-cognitivist externalism would be a combination of the weak form of internalism with externalism about motivation. In fact, what Brink calls “weak internalism” cannot, in my understanding, stand alone, since it can only be part of the non-cognitivist externalist conception of moral motivation. For, in the matter of moral motivation it is my view that we are interested in whether moral judgements have the power to motivate by themselves completely or whether they need for this purpose the assistance of some affective state or attitude. Therefore, even if they provide some motivation, as long as they motivate sufficiently only in conjunction with distinct non-cognitive states (that differ from those expressed in the judgements) we can no longer speak of internalism.} Knowledge is meant to be a presupposition or condition of the moving psychological state in the case of cognitivist but not in the case of the non-cognitivist externalism.

The question we need to address at this point is how Kant is to be placed in this schema.

In effect, Kant had searched at length for an answer to the question concerning how the consciousness of the moral law is connected with a moving force for the performance of the morally good action. In his early Lectures on Ethics he called the answer to this question the “philosopher’s stone”:

The moral feeling is a capacity for being affected by a moral judgment. When I judge by understanding [durch

Konstantinos Sargentis, Moral Motivation in Kant, KSO 2012: 93-121. 
Posted June 12, 2012 www.kantstudiesonline.net
© 2012 Konstantinos Sargentis & Kant Studies Online Ltd.
den Verstand] that the action is morally good, I am still very far from doing this action of which I have so judged. But if this judgment moves me to do the action, that is the moral feeling. Nobody can or ever will comprehend how the understanding should have a motivating power; it can admittedly judge, but to give this judgment power, so that it becomes a motive [incentive / Triebfeder] able to impel the will to performance of an action – to understand this, is the philosopher’s stone [der Stein der Weisen]. (Ethics XXVII, 1428)\(^{19}\)

Kant explains this further by adding:

When man has learned to appraise all actions, he still lacks the motive [incentive / Triebfeder] to perform them. […] The understanding has no elateres animi, albeit it has the power to move, or motiva; but the latter are not able to outweigh the elateres of sensibility. A sensibility in accordance with the motive power of the understanding would be the moral feeling; […] if it [the understanding] could move sensibility to concur and to motivate, that would be the moral feeling. […] When […] sensibility abhors what the understanding considers abhorrent, this is the moral feeling. […] but […] man has no such secret organization, that he can be moved by objective grounds. (ibid, 1429)

The implication in the above passages is that moral judgements not only express beliefs about objective moral facts (cognitivism) but, what is more, in doing so they cannot by themselves move towards action (externalism). However, although Kant introduces the notion of moral

---

\(^{19}\) Cf. CP V, 72: “…how a law can be of itself and immediately a determining ground of the will (though this is what is essential in all morality) is for human reason an insoluble problem”. Due largely to such views, I cannot agree with Dieter Henrich, when, with reference to Kant’s doctrine of the “fact of reason”, he argues in favour of the unity of the knowledge of good and moral motivation in the context of the phenomenon of “moral insight” (see Henrich 1960).
feeling, he constructs his argument by employing the hypothetical ‘if a moral action takes place, then its incentive is the moral feeling’ and he seems unable to show what that feeling consists in. To be sure, he does offer a sort of definition of moral feeling, that is, ‘the concurrence of sensibility and understanding’, yet this definition is too formal to explain the nature and the function of the phenomenon ‘moral feeling’. Thus, Kant here seems to diverge slightly from the externalist requirement, since, in reality, he fails to show what this particular state of mind is which supplies the moving force to the performance of moral action. Therefore, it is my view that his position here may be characterised more properly as cognitivist ‘non-internalism’. However, in any case, he explicitly separates the principle of appraisal from the principle of incentive.\textsuperscript{20} The principium diiudicationis bonitatis is not at the same time the principium executionis bonitatis.\textsuperscript{21} It must be noted that Kant clearly

\textsuperscript{20} Cf. EV XXVII, 274-275: “If the question is: What is morally good or not?, that is the principle of appraisal, whereby I judge the goodness or depravity of actions. But if the question is: What moves me to live according to this law?, that is the principle of motive [incentive / Triebfedér]. Appraisal of the action is the objective ground, but not yet the subjective ground. That which impels me to do the thing, of which understanding tells me that I ought to do it, is the motiva subjective moventia. The supreme principle of all moral judgement lies in the understanding; the supreme principle of the moral impulse to do this thing lies in the heart. This motive [incentive / Triebfedér] is the moral feeling.”

\textsuperscript{21} On this terminology cf. Henrich 1981, 14 ff., and Patzig 1986. Patzig affirms this thesis, when he argues that “it is not enough to have insight into the rightness of a moral norm to also act according to it”, and he goes on saying that “the possible gap between the assent to the validity of a norm and a behaviour which corresponds to that norm belongs to the irrefutable, main facts of human life”, and that “this gap cannot be eliminated even by subtle arguments” (204 f.). According to Patzig, this issue also concerns the moral problem. He maintains that “one can accept a moral norm as relevant and rationally justified without at the same time being prepared to act according to it. The decision to act according to norms is a-rational, or pre-rational and goes back to motives which can concur arbitrarily with other emotional motives” (217). Although Patzig does find this distinction in Kant’s problem of the philosopher’s stone, he, nevertheless, believes – and this is something I do not agree with – that Kant renounces it with his doctrine of the fact of reason and the role of respect in it: “Although the necessary distinction between ‘princīpiūm diiudicationis’ and ‘princīpiūm executionis’ is made in Kant, it is, nevertheless, renounced by the fact that both turn out to be only slightly different aspects of the same thing” (ibid.). The reason for my disagreement with Patzig lies in that here one could quite plausibly ask: Is
implies this separation later in the *Critique of Pure Reason*, when he claims:

It is necessary that our entire course of life be subordinated to moral maxims; but it would at the same time be impossible for this to happen if reason did not connect with the moral law, which is a mere idea, an efficient cause which determines for the conduct in accord with this law an outcome precisely corresponding to our highest ends, whether in this or in another life. Thus without a God and a world that is now not visible to us but is hoped for, the majestic ideas of morality are, to be sure, objects of approbation and admiration but not incentives for resolve and realization (CR A 812 f. / B 840 f.).

In the above passage from the CR Kant introduces the necessary conditions for the possibility of the highest good, i.e., the postulates of God and immortality, as props for the moral law, so that they impart to it “obligating force” (CR A 815 / B 843) and motivating power. At first glance, if one follows Kant’s words literally, it may seem that the hope for the highest good and its conditions is not by itself an incentive, but that it is relevant to motivation only in making the moral law into an incentive. Therefore, if the incentive is the moral law itself, one may think that we have an internalist model of motivation. But this is not true, since here we have neither a non-cognitivist nor a cognitivist type of internalism. For, on the one hand, moral knowledge does not express the affective state of the hope for the highest good, and thus we are not concerned with a non-cognitivist kind of internalism; and, on the other hand, insofar as that

---

Patzig’s “a-” or “pre-rational decision” not ultimately an ‘act’ of the absolutely free or spontaneous will (*Willkür*), as Kant presents it in his ‘Incorporation Thesis’? (cf. n. 27 in this paper). If that is the case, then the above-mentioned distinction is still in place in Kant’s mature moral theory and is not, as Patzig maintains, renounced by Kant.

22 Cf. ibid. A 818 / B 846, where Kant says that we postulate God as a “wise world-regent, in order to give effect” to the moral laws.
hope plays a decisive role in motivation, we cannot have a cognitivist kind of internalism, which entails that the only possibility left open is that of externalism. More specifically, what is involved here is a cognitivist externalist model of motivation; yet, again, it is not clear in which sense it relates to the kind of cognitivist externalism I described in the first part of this paper. As we saw above, following that conception, there must be a certain link between the conative state, the hope for the highest good, and moral knowledge, so that moral knowledge causes or produces the conative state. Here, this is possible only if we take the content of moral knowledge to be morality as “worthiness to be happy” (CP A 806 / B 834). Therefore, it may be said that insofar as moral knowledge expresses the worthiness to be happy, it entails the notion of happiness as part of its content and it causes or activates the affective (but, nevertheless, also rational) state of hope for the highest good (and its conditions, i.e., God and immortality), which, in turn, moves towards action. So, in the CR Kant finds the previously missing external motivating factor of moral action in the affective state of hope for the highest good. In so doing, he fills in the gap in the externalist picture he left open in the earlier Lectures with a positive account of cognitivist externalism.

However, this is not the final form of Kant’s theory of moral motivation. His mature moral theory does not recognise the affective state of hope for the highest good as an incentive for moral action. For, as is often maintained in the literature, after Kant’s introduction of the principle of

23 By saying that the judgement “entails the notion of happiness as part of its content”, I do not mean that it also expresses it, since what the judgement expresses is just worthiness to be happy, which is another name for morality; otherwise we would have a non-cognitivist ethical model, which is clearly not adopted by Kant (even in the first Critique). However, insofar as happiness is mentioned in the definition of morality, it provides us with the needed link between moral judgement and the external affective state.

24 As Klaus Düsing (1971, 27 f) points out, in the CP Kant changes his theory of the highest good, in that he dissociates it from the foundation of Ethics, which was characteristic of the first Critique, and makes it part of the “complete development of the doctrine of the finite ethical consciousness” (28).
autonomy, such recognition would amount to the view that morality is heteronomous. In his major moral works, from the GMM onwards, Kant fills in the above-mentioned gap in his externalist model of the *Lectures* with the moral feeling of respect. In addition, Kant’s theory of respect provides an answer to the question regarding the impact the consciousness of the moral law has on the mind when it motivates, but not to the question regarding how the production of such an effect is possible:

What we shall have to show a priori is, therefore, not the ground from which the moral law in itself supplies an incentive but rather what it effects (or, to put it better, must effect) in the mind insofar as it is an incentive (CP V, 72).

Therefore, on the one hand, Kant still implicitly admits the impossibility of finding out the philosopher’s stone, but, on the other, he now does identify the incentive for moral action within the framework of his autonomous morality.

III

Before we proceed to the specific problem of Kant’s cognitivist externalism, two issues need to be clarified. Firstly, one may wonder whether the fact that respect is an incentive also means that if someone feels respect for the moral law, s/he necessarily performs the right action. In the literature it is often stated that Kant himself has muddied the waters, since he characterises respect as “in itself a sufficient incentive to the power of choice [*Willkür*]”

---

25 See for example Beck 1960, 273 f.; Allison 1990, 67; Höffe 2004, 294, 298; cf. Düsing 1971, 15 f. An exception in the literature is Maximilian Forschner, who shows that Kant has always held the hope for the highest good, which is being given in the rational faith (*Vernunftglaube*), - next to the feeling of respect - an incentive for moral actions (see Forschner 2005). In any case, Forschner’s thesis does not run counter to the externalist picture of Kant’s theory of moral motivation, as I present it here, but rather (as is obvious) supports it.

26 Cf. CP V, 79, where Kant argues that “we cannot cognize…the force of the pure practical law as incentive”.

Konstantinos Sargentis, Moral Motivation in Kant, KSO 2012: 93-121.
Posted June 12, 2012 www.kantstudiesonline.net
© 2012 Konstantinos Sargentis & Kant Studies Online Ltd.
For, if respect is moreover “inseparably connected with the representation of the moral law in every finite rational being” (CP V, 80) and a free action is only an action performed out of respect for the moral law, then it seems that the possibility of an evil or immoral action as a free action is not explainable. Secondly, one may object that the specific problem of motivation does not concern exclusively moral motivation, but rather rational motivation in general. Why does Kant, then, link the question of the “philosopher’s stone” exclusively with moral motivation? How is the moving force of pragmatic or prudential considerations as reasons for action to be explained?

However, those objections disappear as soon as one focuses on Kant’s main philosophical concern. More specifically, Kant’s aim is to show that pure practical reason can in itself motivate towards action; that beside the principium divinationis bonitatis it also produces a principium executionis bonitatis, which can be effective without the cooperation of (and even contrary to) other incentives; that it can really be ‘autonomous’. For this conception of autonomy, the required principium executionis is respect for the moral law, which is inseparably connected with the representation of the moral law in every finite rational being. (Rel VI, 27).

27 For example Harald Köhl (1990, 139) maintains that due to the phenomenon of the weakness of the will, which is also recognised by Kant, Kant’s thesis that respect is of itself a sufficient incentive to the will [Willkür], cannot be right. I believe that this critique is wrong, because respect is a sufficient incentive in the sense that it alone, i.e. without the cooperation of non-moral incentives, is able to determine the will (Willkür) towards the performance of an action. In my view, the crucial point in Köhl’s critique lies in the fact that respect cannot be an incentive automatically. But, this is consistent with Kant’s moral psychology, as the so-called “Incorporation Thesis” (see Allison 1990, 6, 38 ff.) shows. More specifically, according to the Incorporation Thesis, “freedom of the power of choice has the characteristic, entirely peculiar to it, that it cannot be determined to action through any incentive except so far as the human being has incorporated it into his maxim (has made it into a universal rule for himself, according to which he wills to conduct himself); only in this way can an incentive, whatever it may be, coexist with the absolute spontaneity of the power of choice (of freedom).” (Rel VI, 23 f.)

28 Cf. Rel VI, 27 f.

29 Cf. GMM IV, 447: “...a free will and a will under moral laws are one and the same. If, therefore, freedom of the will is presupposed, morality together with its principle follows from it by mere analysis of its concept”; cf. also CP V, 29, 31, 93-94. Allison (1990, 201 ff.) has called this thesis the “Reciprocity Thesis.”
law. Yet, whether the autonomous agent is, in turn, free to place the moral incentive over the incentives of sensibility or not, is something that does not affect the specific problem of motivation through respect. To be sure, the moral subject does possess that kind of freedom, since next to his/her autonomy s/he has the freedom of a ‘spontaneous’ will (Willkür). What that notion of freedom shows is that in Kant there is not an alleged coercion of reason. Yet, this notion of freedom is not part of Kant’s conception of autonomy. Hence, its analysis does not belong to Kant’s main philosophical program of grounding morality, and is thus not connected to the problem of moral motivation. Similarly, the objection concerning how rational motivation through reasons in general is possible refers to a question that does not arise from Kant’s main concern, which is moral motivation per se and not motivation in general. In sum, Kant discusses the problem of motivation only in the context of his main aim, which is to show “that there is pure practical reason” or that “pure reason […] is really practical” (CP V, 3). Therefore, the real problem in Kant’s theory of moral motivation relates neither to the possibility of immoral action nor to the problem of rational motivation in general.

Nonetheless, the problem of rational motivation in general is one which Kant has tackled, albeit rather tacitly: It amounts to the problem of the possibility of hypothetical imperatives, as Kant articulates it in the GMM. In referring to both the imperatives of skill and the imperatives of prudence, Kant says:

[S]ince both merely command the means to what it is presupposed one wills as an end, the imperative that commands volition of the means for him who wills the end is in both cases analytic. Hence there is also no difficulty with respect to the possibility of such an imperative. (GMM IV, 419).30

30 For the problems concerning the possibility of hypothetical imperatives, see Timmermann 2007, p. 69 ff.

Konstantinos Sargentis, Moral Motivation in Kant, KSO 2012: 93-121.
Posted June 12, 2012 www.kantstudiesonline.net © 2012 Konstantinos Sargentis & Kant Studies Online Ltd.
Interestingly enough, the problem of the philosopher’s stone arises again when Kant turns to the problem of the possibility of the *categorical* imperative. In particular, he proceeds to tackle it in a reductive manner, that is, by citing the categorical imperative’s presupposition, which is the idea of freedom, and he argues that we cannot prove the possibility of freedom *itself*. Without stating it explicitly, he equates this latter impossibility with the philosopher’s stone. As he says:

[I]t is quite beyond the capacity of any human reason to explain how pure reason, without other incentives that might be taken from elsewhere, can be of itself practical, that is, how the mere *principle of the universal validity of all its maxims as laws*...can of itself furnish an incentive and produce an interest that would be called purely *moral*; it is impossible for us to explain, in other words, *how pure reason can be practical*, and all the pains and labor of seeking an explanation of it are lost. It is just the same as if I tried to fathom how freedom itself as the causality of a will is possible. For then I leave the philosophic ground of explanation behind and I have no other. (GMM IV, 461 f.)

In consequence, the problem of moral motivation in Kant is exclusively about the nature and the function of respect as a product of pure reason and as an incentive for moral actions. In the subsequent two sections of this paper, I will further explore the nature and the function of respect.

## IV

In order to examine the nature of respect one has to consider the essential features of respect, which can be extracted mainly from the GMM and the CP. Respect is

---

31 Cf. CP V, 72: “…how a law can be of itself and immediately a determining ground of the will (though this is what is essential in all morality) is for human reason an insoluble problem *and identical with that of how a free will is possible*” [emphasis added].

Konstantinos Sargentis, Moral Motivation in Kant, KSO 2012: 93-121.

Posted June 12, 2012 www.kantstudiesonline.net

© 2012 Konstantinos Sargentis & Kant Studies Online Ltd.
a feeling, yet it is a feeling of a special kind:

    [T]hough respect is a feeling, it is not one received by means of influence; it is, instead, a feeling self-wrought by means of a rational concept and therefore specifically different from all feelings of the first kind, which can be reduced to inclination or fear (GMM IV, 401 n.).

In the CP (V, 73) Kant states that respect is a feeling produced by an “intellectual ground”, and, what is more, that we know it “a priori”: “[…] respect for the moral law is a feeling that is produced by an intellectual ground, and this feeling is the only one that we can cognize completely a priori and the necessity of which we can have insight into.” Moreover, we are further told that respect has “an intelligible cause” (ibid., 75). Thus, Kant makes it clear that respect differs from other feelings in that it has an intellectual side. However, it is important to ask what this intellectual side consists in.

To answer the question we may begin by arguing that the intellectual side of respect primarily consists in its relation to the moral law and to the law-giving activity of pure practical reason. Respect is “the effect of consciousness of the moral law”, and its intelligible cause is “the subject of pure practical reason as the supreme lawgiver” (CP V, 75). “Immediate determination of the will by means of the law and consciousness of this is called respect, so that this is regarded as the effect of the law on the subject” (GMM IV, 401 n.). Unless we conceive the notion of “effect” here as an ‘act’ performed by the human will, the above-cited portrayal of respect as “self-wrought” cannot make sense. 32 This ‘act’ is performed by the autonomous will which is the “proper

---

32 Although Kant implies this ‘solution’ both in the CP and in the GMM, his words are, nevertheless, obscure. Cf. GMM IV, 400: “For an object as the effect of my proposed action I can indeed have inclination but never respect, just because it is merely an effect and not an activity of a will” [last emphasis added]; and CP V, 79: “Recognition of the moral law is … consciousness of an activity of practical reason from objective grounds” [emphasis added].
self" of man. In order to explain in which sense the ‘effect’ of the moral law is an ‘act’ of the will, we need Kant’s later distinction between *Wille* and *Willkür*, which amounts to a distinction between the legislative and the executive function of the (autonomous) will as a unified faculty. Given that distinction, it may be said that the will as *Wille* produces by itself the moral law and imposes it upon itself as *Willkür*. At the same time, the will as *Willkür* assents to the moral law and approves it, so that respect is the recognition of the law as an objective worth:

Respect is properly the representation of a worth that infringes upon my self-love. Hence there is something that is regarded as an object neither of inclination nor of fear, though it has something analogous to both. The object of respect is therefore simply the law, and indeed the law that we impose upon ourselves and yet as necessary in itself. …As imposed upon us by ourselves it is…a result of our will (GMM IV, 401 n.).

What is referred to above as ‘worth’ (*Wert*) is grounded on the fact that the moral law is a product of the rational nature of man and of the autonomy of his will.

So far we have examined the intellectual side of respect. But, due to its relation to sensibility, respect also has an affective side. In any case, respect is a feeling and as such it is a defining feature of finitude. Moreover, although respect is a feeling with an intellectual ground, it still arises

---

33 On the notion of “proper self”, see GMM IV, 458, 461.
34 See MS VI, 226.
35 Concerning the act of the approval of the moral law as well as the assent to it within the framework of the phenomenon of “moral insight” in Kant’s doctrine of the fact of reason, see Henrich 1960, 83-88.
36 Cf. Köhl 1990, 132. In this context Köhl does not use the term “worth”, but “dignity” (which is attributed to reason as a cause of respect).
37 See CP V, 76: “[…] it should be noted that, since respect is an effect on feeling and hence on the sensibility of a rational being, it presumes this sensibility and so too the finitude of such beings on whom the moral law imposes respect, and that respect for the law cannot be attributed to a supreme being or even to one free from all sensibility, in whom this cannot be an obstacle to practical reason.”
from the impact this intellectual ground has on sensibility. It primarily relates to the negative effect on the inclinations of the moral law as a determining principle of the will:

For, all inclination and every sensible impulse is based on feeling, and the negative effect on feeling (by the infringement upon the inclinations that takes place) is itself feeling. Hence we can see a priori that the moral law, as the determining ground of the will, must by thwarting all our inclinations produce a feeling that can be called pain; and here we have the first and perhaps the only case in which we can determine a priori from concepts the relation of a cognition (here the cognition of a pure practical reason) to the feeling of pleasure or displeasure. (CP V, 72 f.).

Kant explains further that the infringement of inclinations is a weakening and a ‘breaking down’ or a humiliation of the human relation to sensibility, especially when that relation takes the form of self-conceit:

All the inclinations together (which can be brought into a tolerable system and the satisfaction of which is then called one’s own happiness) constitute regard for oneself (solipsismus). This is either the self-regard of love for oneself, a predominant benevolence toward oneself (Philautia), or that of satisfaction with oneself (Arrogantia). The former is called, in particular, self-love; the latter, self-conceit. Pure practical reason merely infringes upon self-love, inasmuch as it only restricts it, as natural and active in us even prior to the moral law, to the condition of agreement with this law, and then it is called rational self-love. But it strikes down self-conceit altogether, since all claims to esteem for oneself that precede accord with the moral law are null and quite unwarranted (CP V, 73).

What Kant has in mind is specifically this negative effect on sensibility, when he equates respect with “submission to a
law” (ibid., 80) or with the consciousness of the subordination under the moral law:

What I cognize immediately as a law for me I cognize with respect, which signifies merely consciousness of the subordination of my will to a law (GMM IV, 401 n.).

Now, since what strikes-down or humiliates is the moral law as “the form of an intellectual causality, that is, of freedom” (CP V, 73), which is something “positive in itself”, the moral law also generates respect for itself. “The humiliation on the sensible side, is an elevation of the moral, i.e., practical, esteem for the law itself on the intellectual side” (CP V, 79). Indeed, respect is a feeling inasmuch as it cannot arise without the negative effect of the moral law on sensibility (as in the case of a pure rational being). However, it may be called a feeling even in its ‘elevating’ relation to the moral law, since this positive relation to pure practical reason depends on the negative aspect of the generation of respect. Kant maintains that the negative effect on sensibility is an “out-moving of a resistance” or a “removal of a hindrance”, which in the “judgment of reason” counts as a “positive furthering of its causality”; and that “on both grounds together” this feeling “can be called moral feeling” (CP V, 75).

To sum up, respect is a complex phenomenon. It has an intellectual side, which is indicated by such notions as “self-wrought”, “effect of the law on the subject”, “representation of a worth”, “intellectual ground”, “intelligible cause”, and “elevation of the practical esteem for the law”. But, it also has an affective side, which is indicated by notions, such as “feeling”, “sensibility”, “infringement upon the inclinations”, “humiliation”, and “finitude”. Of course, although the affective side of respect is primarily negative, we should not identify it as exclusively negative. Insofar as respect in its elevating relation to the moral law and in its furthering the law’s causality is a feeling, the affective side of respect is also to be seen as positive. Following the examination of the
nature of respect, I will now turn to the function of respect as the incentive for moral action.

V

In the previous part of this paper we examined the basic features of respect as a phenomenon which belongs to every human being’s autonomous will. However, Kant introduces the feeling of respect in the first place not to specify its place and its role within the autonomous constitution of the will, but rather to explain the possibility of moral action:

This feeling (under the name of moral feeling) is [...] produced solely by reason. It does not serve for appraising actions and certainly not for grounding the objective moral law itself, but only as an incentive to make this law its maxim. (CP V, 76).

The question we need to pose at this stage concerns the particular factor which supplies the moving force towards action. More specifically, is this factor the intellectual side or the affective side of respect (both of which were considered in the previous section)?

Kantian scholars are divided into two groups depending on the view they adopt with reference to the latter question; that is, they are either intellectualists or affectivists. Richard McCarty, who is an affectivist, explains this distinction in the following way:

Intellectualists hold that respect for the moral law is, or arises from, a purely intellectual recognition of the supreme authority of the moral law, and that this intellectual recognition is sufficient to generate moral action independently of any special motivating feelings or affections. [...] Affectivists need not deny that Kantian moral motivation initially arises from an intellectual recognition of the moral law. Contrary to intellectualists, however, they
maintain that it also depends on a peculiar moral feeling of respect for law, one consequent to the initial recognition of moral judgment the intellectualists emphasize exclusively.\footnote{McCarty 1993, 423.}

Following the intellectualist reading, the affective side of respect is merely a result of the process of motivation. For instance, Andrews Reath maintains that respect as a feeling presupposes the check of the inclinations. As a result, he comes to the conclusion that the motivational power of respect as an incentive may be found only in the intellectual aspect of the phenomenon termed “respect” (as opposed to the affective aspect, i.e., respect as a feeling).\footnote{See Reath 1989, 286 ff.} My main disagreement with the intellectualist reading concerns the fact that I believe – and this is often overlooked in the literature\footnote{An exception is McCarty, who discerns in Kant two phases in the process of moral motivation, a ‘cognitive’ and an ‘affective’ one (see McCarty 1994, 19, 23).} – that in the CP Kant has two conceptions of “incentive”. At the beginning of the motivational process, it is the moral law itself which is an incentive towards the determination of the will (but not yet towards the action), so that respect with its affective aspect arises:

So far, then, the effect of the moral law as incentive is only negative, and as such this incentive can be cognized a priori. For, all inclination and every sensible impulse is based on feeling, and the negative effect on feeling (by the infringement upon the inclinations that take place) is itself feeling [emphasis added] (CP V, 72 f.).

If we employ here the distinction Kant makes in the GMM between a “motive” (Beweggrund) and an “incentive” (Triebfeder)\footnote{See GMM IV, 427: “the subjective ground of desire is an incentive; the objective ground of volition is a motive”.}, it seems that the moral law is a motive, rather than an ‘incentive’, and what it ‘motivates’ through its nega-
The effect on inclinations is the genesis of respect as a feeling. This respect is, in fact, as Reath maintains, “an emotion that is the effect of, and follows from, the determination of the will by the Moral Law, when the latter limits the inclinations”\(^{42}\). But, in its turn, the feeling of respect is in itself also an incentive (rather than a ‘motive’) towards moral actions:

This feeling…does…serve…only as an incentive to make this law itself its maxim.” (CP V, 76); “…respect for the moral law must be regarded…as a subjective ground of activity – that is, as the incentive to compliance with the law – and as the ground for maxims of a course of life in conformity with it (ibid., 79).

Consequently, it seems that the process of moral motivation in Kant consists of two main phases. (1) In the first main phase, the moral law ‘motivates’ the generation of respect as a moral feeling. In its turn, this phase consists of two sub-phases, a negative and a positive one. (1a) What takes place in the negative sub-phase, is the negative effect of the moral law on the inclinations, that is, the infringement upon self-love and the humiliation of self-conceit, the ‘out moving of a resistance’ or the ‘removal of a hindrance’. Since “the negative effect on the feeling is itself feeling”\(^{43}\), what arises in this sub-phase is respect as a mere feeling.\(^{43}\)

Now, at the same time this negative effect on feeling counts “in the judgement of reason” as something positive, which is what the positive sub-phase consists of. (1b) In the positive sub-phase, what takes place is the ‘elevation of the practical esteem for the law’ and the ‘positive furthering of its causality’. What arises here is not respect as a ‘mere’ feeling, but

---

\(^{42}\) Reath 1989, 287

\(^{43}\) It should be noted that although the negative effect upon feeling in general is “pathological” (CP V, 75), in the case of the feeling of respect the negative effect of the moral law on sensibility cannot be considered as pathological: “this feeling, on account of its origin, cannot be called pathologically effected but must be called practically effected” (ibid.), since its ground is not “an object of the senses” (ibid., 80), but the moral law as a product of pure practical reason.
“a feeling of respect for the moral law” (ibid., 75). Moreover, what arises from both sub-phases taken together should be called neither a ‘mere’ feeling nor a ‘feeling of respect for the moral law’, but rather a ‘moral feeling’, as Kant emphasises in CP V, 75. In addition, because these two sub-phases take part only in the ‘judgement of reason’, we should conceive them as two different aspects of the same thing, that is, of the effect of the moral law on sensibility.\(^{44}\)

(2) In the second main phase, the ‘moral feeling’, which is produced in the first main phase, serves as an incentive for the agent to make the moral law his/her maxim and to perform moral actions.

As in the case of the two sub-phases of the first main phase, the two main phases may be considered from the standpoint of the “judgement of reason” as two aspects, an objective and a subjective one, of a unified process termed ‘moral motivation’. Regarded objectively, a moral action is one which is determined by pure practical reason; regarded subjectively, it is performed through the moving force of respect as a feeling and as an incentive, whilst this feeling is not an “antecedent feeling in the subject that would be attuned to morality” (ibid.); rather, it is “practically effected” (ibid.). It is in this sense of the objective and the subjective aspect of a unified process, as it seems to me, that Kant states in the CP (V, 76):

\(^{44}\) It must be acknowledged that it is not easy to show in which sense the two sides of respect are two different aspects of the unified phenomenon of respect. In the literature, this problem has led to a confrontation between Henry Allison and Dieter Henrich. On the one hand, Henrich (1981, 36 ff.), under Hegelian influence, emphasises Kant’s incapability of accounting for respect as a unified phenomenon of ethical consciousness. On the other hand, Allison (1990, 127 ff.) holds this critique to be mistaken, since, according to his view, Henrich overlooks the role and the importance of the “judgment of reason” in Kant’s theory of moral motivation. However, Melissa Zinkin (2006) has recently suggested an interesting solution to this problem. By referring to Kant’s pre-critical essay, Attempt to Introduce the Concept of Negative Magnitudes into Philosophy (1763), Zinkin attempts to show that “Kant’s account of how the moral law effects in us a feeling of respect is underpinned by his view that the will is a kind of negative magnitude, or force” (33). By means of this argumentation Zinkin aims to show that respect is a feeling which is, nevertheless, not separate from the use of reason.
And so respect for the law is not the incentive to morality; instead it is morality itself subjectively considered as an incentive.

In this puzzling dictum Kant wants to make clear that the feeling of respect is produced by the moral law, and that it is the means the moral law uses in order to determine the will towards action. Therefore, he points out that insofar as the feeling of respect is produced by the moral law, it may be said that it is “not the incentive to morality”, because if it were, it would be an “antecedent feeling in the subject that would be attuned to morality”; instead, as a product of the moral law, it is nothing but “morality itself”, and it is ‘subjectively’ an incentive, if we take into account that the ‘objective’ determining ground is pure practical reason. Thus, we can see the two phases of moral motivation as two sides of the same coin. In any case, the fact that the second main phase, or aspect, of the process of moral motivation in Kant takes incentive for moral actions to be respect as feeling indicates that the affectivist reading of the theory of respect is the most appropriate one.45

VI

What is implied in the affectivist reading is the assumption that an “additional psychological sanction” (in Nagel’s language) and not moral knowledge as such is what constitutes the incentive of, or supplies the motivation for, moral action. This renders Kant an externalist concerning moral motivation. Nonetheless, as we saw

45 It is for this reason that I agree with Jeanine Grenberg when she recognises the motivating role of feeling in moral action, and, what is more important, when she underlines the impossibility of motivation without the contribution of the affective, sensible aspect of incentive in “Kant’s Theory of Action”: “Maxims cannot, in themselves, be a sufficient motivation to action for finite rational beings; [...] The act of incorporation [of drives into maxims – K.S.] thus also, beyond the generation of a principle of action, involves the acquisition of the force of one’s sensible drives in the service of the agent’s rational choice” (Grenberg 2001, 178). An interpretation of the Kantian problem of moral motivation and of the theory of respect, which I agree with in most respects, is that of McCarty’s (1993; 1994).
above, moral knowledge in Kant (as presented in the intellectual side of respect) plays a decisive role in the genesis of respect as a feeling and as an incentive, and thus, it is pertinent to motivation. This role of knowledge, insofar as knowledge expresses beliefs about moral facts, renders, in its turn, Kant’s externalism a *cognitivist* kind of externalism.46

One may wonder why I have placed such emphasis on the distinction between the cognitivist and the non-cognitivist versions of internalism and externalism, since Kant’s externalism alone, as contrasted to the standard internalist interpretation of his theory of moral motivation, seems to have been my basic aim in this paper. Yet, I consider it important to emphasise Kant’s cognitivism, in that the reason the advocates of the usual internalist interpretation of Kant employ for arguing that Kant is an internalist, is precisely the reason which makes Kant a *cognitivist* externalist.47 And this is so, because, despite moral knowledge

46 Following Brink’s distinctions between (a) *agent*, (b) *appraiser*, and (c) *hybrid* motivational models (see Brink 1989, 40-43), one could here further ask what we mean when we say that the moral knowledge is pertinent to motivation. Is the motivationally relevant factor (a, objectively) the *moral fact* or the *moral obligation*, regardless of the appraiser’s recognition of it, (b, subjectively) the *belief* about the moral fact, regardless of whether this belief is correct, or (c, both objectively and subjectively) *the recognition of the moral fact* on the part of the agent? According to these distinctions, the Kantian version of externalism, as I present it here, would be a “hybrid” motivational model.

47 I find worth mentioning Ralph Walker’s examination of respect in the GMM (see Walker 1989). Although Walker gives the impression that Kant is an internalist, I do not believe that behind his interpretation lies a distinction between internalism and externalism (with the result that Kant is taken to be an internalist). Rather, he relies on a distinction between cognitivist and non-cognitivist externalism. The upshot is that, according to my view, in Walker’s paper Kant should be interpreted as a cognitivist externalist. Cf. two crucial places in Walker 1989: a. (p. 98) “The passages in the *Grundlegung* about moral interest and about *Achtung* as feeling imply that [...] it can motivate us only by mysteriously creating a special kind of feeling which then pushes us to act” (i.e., Kant is an externalist ‘about motivation’); b. (p. 105) – where Walker *opposes* the following view of Mackie: “…the moral law is not objective, not something that obtains independent of human moral sentiments and beliefs, but is rather simply the projection of attitudes that human beings naturally have” (i.e., according to Walker Kant is not a non-cognitivist). Thus, I believe that in Walker’s paper Kant in the GMM is presented as a cognitivist externalist.
being necessary for the genesis of the feeling, the moving force towards action stems exclusively from the feeling as such and is independent from moral knowledge: it is ‘external’ to it. Of course, one may argue that insofar as the feeling is produced by knowledge, to a certain extent knowledge and feeling cooperate in bringing about action. But again, if we want to identify which of the two is the source of the moving force towards action, then we have to ascribe this property exclusively to feeling.

However, it must be noted that the above-called ‘independence’ of the moving force of moral knowledge is not meant in the sense assumed by the ‘standard’ externalist view. As was shown above, respect is a peculiar feeling with an intellectual basis, and is related negatively to sensibility and to the needs of finite ‘man’: that is, insofar as the moral law limits the inclinations. But there is no positive relation between respect as feeling and the needs of man as a sensible being: and this is the main divergence of Kant from the standard externalist picture. The peculiarity of the moral feeling of respect is at the same time the reason why Kant’s externalist model of moral motivation is peculiar. Nonetheless, this peculiarity is not a reason for denying that Kant is an externalist.

Note

References to Kant’s works are by volume and page number of the Akademie Edition (1902 ff.). References to the Critique of Pure Reason are to the first and second editions. The English translations are from the ‘Cambridge Edition of the Works of Immanuel Kant, general editors: Paul Guyer and Allen Wood, Cambridge University Press’. Abbreviations: CP (Critique of Practical Reason), CR (Critique of Pure Reason), Ethics (Lectures on Ethics), GMM (Groundwork of the Metaphysics of Morals), MM (The Metaphysics of Morals), Rel (Religion within the Boundaries of mere Reason). – The translations from the German secondary literature are the author’s.
Bibliography


