

Received November 22nd, 2007

**Sonja Rinofner-Kreidl**

Universität Graz, Institut für Philosophie, Heinrichstraße 26, A-8010 Graz  
sonja.rinofner@uni-graz.at

## **Compatibilistic Visions**

### **A Response to Michael Pauen's "Self-Determination, Free Will, Responsibility, and Determinism"**

Michael Pauen defends the compatibility of freedom and determinism by way of strengthening the principle of authorship and interpreting the principle of alternative possibilities in terms of determinism. Authorship is said to be incompatible with indeterminism because the latter is unable to grasp the connection between the mental content of an agent (desires, beliefs and so on) and her action in a non-fortuitous way. Apart from authorship, there is a second minimal criterion which, according to our common sense view of freedom, must be met, namely autonomy. Authorship requires absence of pure chance, autonomy requires absence of compulsion. In course of specifying his idea of self-determination, the author examines current arguments against compatibilism. The paper sets in with a terse introduction which nicely formulates the issue and, among other things, rightly stresses that every theory of self-determination depends on some idea of the self which has to be made explicit.

Although the author's statements are mostly clear and unfolded in a consistent way, the whole project is burdened with serious difficulties. The latter are due to the basic approach Michael Pauen follows in reconsidering the issue of self-determination. Thus, it is not surprising to find these difficulties manifesting themselves in the guiding concepts as well as in the type and range of questions the author takes into consideration. Consequently, we cannot restrict ourselves to asking whether the author's argumentation is consistent with regard to a given framework of presuppositions. Instead, we shall have to scrutinize whether the compatibilistic approach is suited to do justice to the ideas of self-determination and freedom of will.

From a compatibilistic point of view it is near at hand to formulate the issue of freedom in terms of explaining actions with reference to given preferences. Accordingly, the true question seems to be what we should think about the connection between rationality or, to give a more precise idea, rational choice, on the one hand, and freedom, on the other hand. It is in this context that Pauen argues that an action may be considered self-determined if (and only if) it follows from my own desires and beliefs. An action fails to be self-determined if this is not the case, i. e. if there either lacks a relevant connection between these mental contents and my present way of acting or the latter occurs by way of compulsion. In the first case the relevant behaviour is indistinguishable from a fortuitous occurrence and, hence, could not be regarded

as someone's action at all. In the second case we are confronted with some kind of involuntary behaviour. Consequently, an action may be said to be self-determined if (and only if) it is determined by my own preferences (see "determination by the agent's preferences is everything we need for a self-determined choice", p. 21\*). According to the author, self-determination in this sense is entirely indifferent from the question of how the preferences at issue came into being (p. 12f). Moreover, we take it that it is only a rational agent whose acting can be self-determined (see p. 6f). This, of course, does not mean that we should consider every action realized by rational agents to be self-determined.

In order to scrutinize Michael Pauen's theory of self-determination, one carefully has to take notice of the concept of personal preference which plays a vital role in his reasoning. According to Pauen's so-called liberal account of determining preferences (p. 8ff), something can be called "personal preference" only on condition that it shows a certain temporal stability (pp. 7, 12) and, still more important, is a possible object of self-determined decisions. Consequently, identifying personal preferences we suppose that the items in question (beliefs, attitudes and so on) are, on principle, subject to self-control, i. e. to a rationally guided process of changes. Introducing the concept of personal preference in this way allows the author to maintain that irrational tendencies could not be reflected in the formation of preferences. Therefore, psychical as well as physiological addictions, for instance, can be classified as non-personal preferences since the relevant kinds of behaviour elude every attempt to achieve self-control.<sup>1</sup> At least partly this approach seems to meet a common sense view on addiction. Normally we take it that although it may occur that I succumb to seemingly irresistible desires or sensual appetites at a certain moment, such transient motives and scattered instances of addictive behaviour do not give rise to an overall, long-term idea of an addicted self. Preferences constitute constant patterns within the self-model of a person which represent a more or less strong bias in favour of rationality (see p. 20). Due to this concept of personal preference the author is able to unhinge the argument that, in case that an addict consents to his addictive behaviour, we should feel compelled to consider his actions as self-determined. Pauen can refuse this argument since, according to his idea of self-determination, addictive behaviour *per definitionem* cannot constitute preferences.

Given that "self-determination" means that an action is determined by personal preferences, one will be eager to hear how this approach can be reconciled with the principle of alternative possibilities. According to this principle, a decision or action is free if the agent could have acted otherwise *under the relevant circumstances*. In this context it becomes evident that Pauen's attempt to reconcile freedom and determinism is realized at a heavy price. The relevant difficulties can be expounded in a twofold manner. First, compatibilism requires a serious weakening of the concept of freedom. Secondly, it strongly endorses an idea of the self which is incompatible with our common sense idea of acting persons. This is noteworthy since the author, in general, makes a point of keeping our "natural" views on freedom in the course of expounding a compatibilistic view of self-determination. Michael Pauen proposes to interpret the principle of alternative possibilities as follows: If I had other preferences (in a given situation), I would have acted otherwise. However, this hardly can be meant as meeting the above-mentioned condition "under the relevant circumstances" since changes of preference certainly give rise to changes of relevant circumstances. For instance, we should expect the agent

to recognize other possible alternatives in case that her preferences have undergone some considerable change. Equally, we expect the range of alternatives as a whole to be widened or restricted in this case. Following Pauen's approach we, moreover, have to argue that, *given* the personal preferences of the agent and *given* some relevant circumstances (either changed or unchanged with regard to some former state), there is only one way of acting which can ensue. This is true, at least, if we suppose the agent to behave rationally. (Nowhere does the author try to query this supposition of rational behaviour. Compare what we said above about his way of introducing the concept of personal preference.) The relevant outcome of Pauen's reasoning may be captured as follows: Self-determined actions result from internal determinations, i.e. determinations by means of personal preferences. Our decisions and actions are free due to the fact that preferences can be ascribed and actions can be explained referring to given preferences. Yet, it is not at all plausible to argue that this reconstruction could pass for an adequate representation of our idea of free will.

Beyond that we may doubt whether the idea of self corresponding to the author's argument is consistent and, in terms of life-worldly experience, sufficiently complex or rich in order to represent our idea of free action. From the above it follows: "So if you know the situation and if you know the agent's preferences then you should know what his choice will be" (p. 21). Forebearing how a particular person will act under specific circumstances is only possible if constant preferences are ascribed. In other words: We shall not be able to successfully forebear a person's actions unless we exclude changes of preference. On the other hand, the author refers to changes of preference in order to defend compatibilism with a view to the principle of alternative possibilities. In other words: Actions cannot be considered free unless we assume possible changes of preference. We are confronted here with a peculiar tension between the idea of freedom and the idea of explaining actions. Within a compatibilistic framework this tension can only be dissolved by reformulating the problem which produces this very tension. Consequently, we should interpret the issue of freedom in terms of whether or not we may succeed in forebearing future changes of preferences. Actually, this kind of reformulation which directs our attention to the deterministic or non-deterministic conception of the self is adverse to an explicit intention of the author: "the freedom of a person seems to result from the freedom of the action she performs – not the other way round" (p. 3). Am I free to become another person as I am now? If this is the case, how can a change of this kind be understood thoroughly? If this is not the case, how could we warrant free will with regard to particular actions?

How is it possible, in accordance with Pauen's approach, to implement freedom in terms of changes of preferences? If we take seriously the compatibilistic approach, we should argue as follows. Changes of preferences either occur by chance or they are determined. Given that fortuitous occurrences cannot be ascribed to agents, and ascribability is a necessary condition of free actions, we should take changes of preferences to be realized in a deterministic way.

\*

Pages in parentheses are related to file-version of Pauen's text (editor).

1

Here, I cannot dwell on the issue whether this really represents a promising and plausible

idea of addiction. For an alternative view see R. Jay Wallace, "Addiction as a Defect of the Will", in: R.J.W., *Normativity and the Will: Selected Essays in Moral Psychology and Practical Reason*, Oxford University Press, Oxford 2006, pp. 165–189.

This is the view Pauen defends, too. Nevertheless, the arguments offered in favour of this view do not bear closer examination. With regard to self-determined actions we can make more or less high demands in terms of rationality. Making heavy demands we may argue that it is not only our decisions with regard to a variety of actually given preferences which should be regulated by means of rational principles. The formation of preferences should be amenable to reason, too. Pauen rejects this stronger demand for rationality because

“... this view [...] would lead us to conclude that one never acts in a self-determined manner if one acts irrationally. Again, given that moral principles can be rationally justified, we had to conclude that no one acts in a self-determined manner if he violates moral principles and nobody would be responsible for immoral acts.” (p. 8)

Do we have to agree to this view? Firstly, we should remember that it was the author himself who stressed the fact that one should not assume every action of a rational agent to be self-determined. Secondly, the idea that a rational formation of preferences, especially a formation of preferences due to rationally justified moral principles, necessarily means that nobody could be called to account for irrational actions is a sound idea only on condition that we live in a compatibilistically interpreted universe. On other conditions, it is not plausible at all. Take, for instance, Aristotle’s idea of practical reason which does not focus on particular actions and their (more or less clearly stamped) rational qualities. Instead, Aristotle endorses the idea of a good life with regard to the temporal and developmental perspectives of a self-determined agent who is expected to assess her preferences, deliberations, processes of decision-making and acting according to the idea of a gradually evolving personal form of life and moral personality, respectively. From this point of view it is not true that demanding for a rational formation of preferences necessarily results in irresponsibility with regard to irrational (or: immoral) actions. This is evident, for instance, when Aristotle, in book III of his *Nicomachean Ethics*, talks about responsibility with regard to one’s own dispositions, thereby anticipating a central aspect of the famous discussion on *akrasia* which is unfolded in book VII afterwards. Contrary to Pauen’s view and in accordance with Aristotle, we may hold agents responsible for their irrational actions *even if these actions are not self-determined*. Taking this view requires to endorse a conception of the self according to which the practical identity of persons is based on their ability to enter into long-term engagements and long-term projects of (self-)education. (The latter, of course, include our rational abilities.) We need not defend the idea of a deterministic change of preferences *because* we otherwise had to assume that nobody could be called to account for irrational actions. This sketchy side-glance to an alternative idea of agency should have made clear that our attempts to acutely describe self-determined decisions and actions finally depend on our idea of the self. It is clearly this idea which is the pivotal point of our debates on free action.

What does that mean with regard to the present context? Since the author’s idea of self directly results from given preferences including their (deterministically grasped) changes which occur in course of time, we are faced with a deterministic conception of the self. This conception should be challenged. According to the compatibilistic view presented above, the self is nothing but a pragmatically isolated part of a comprehensive network of interwoven chains of causes and effects. The rational agent represents a specific interpretation of this network or a specific approach to parts of this network. The rational agent is nothing but a segment of chains of causes and effects which allows to produce a great number of reliable forecasts (due to a complete set

of relevant hypothetically formulated laws of nature and due to the lack of special conditions eluding experimental control). If our talk about the “self” is empty in terms of intentions, purposes, moral obligations and the like, there is no self-determination either since the latter is said to exclude external determination in favour of internal determination, i.e. “determination by the self whose action it is” (p. 5). Therefore, we may conclude that the strongest argument against the compatibilistic approach discussed above has been formulated by the author himself, though unintentionally, when he pointed out that it “is completely unclear what it means to act in a self-determined way, as long as it remains to be spelled out what the self is” (p. 5).