

# Editorial

Die Herausgeber des Philosophischen Jahrbuchs sind in der Regel um eine Kombination von thematischer Konzentration und Vielfalt bemüht. Diesmal überwiegt die Vielfalt – sowohl in historischer wie auch in systematischer Perspektive: Der Beitrag von Thomas Buchheim (München) befasst sich mit der Frage, wie insbesondere Schelling durch seinen Begriff der intelligiblen Tat Einwänden gegen die Möglichkeit freier Selbstbestimmung entgegengehen kann. Das bietet einen aussichtsreichen Ausgangspunkt, um in der aktuellen Debatte etwa Galen Strawsons Position entgegenzutreten. Der Beitrag von Steffi Schadow (Bonn) setzt sich kritisch mit Elisabeth Anscombes These auseinander, wonach die Begriffe der moralischen Pflicht und des Sollens aufgegeben werden sollten. Martina Roesner (Wien) untersucht in ihrem Beitrag das Verhältnis von Philosophie und Medizin am Beispiel von Avicenna und Nietzsche.

Oliver Victor (Düsseldorf) diskutiert mit Blick auf die Philosophie Albert Camus' die Frage, inwiefern die Philosophie popularisiert werden kann und soll. Elena Corsi (Berlin) geht in ihrem Diskussionsbeitrag dem philosophischen Verhältnis zwischen Hans Cornelius und Theodor W. Adorno nach und trägt damit zur Erhellung des Verhältnisses der Kritischen Theorie zum Neukantianismus und zum Neopositivismus bei. Klaus Kienzler (Augsburg) diskutiert Bernhard Caspers Interpretationen von Emmanuel Levinas' „Aufzeichnungen aus der Gefangenschaft“.

Mit dem Aufsatz „Freiheit. Ein Versuch Gott zu denken“ von Hermann Krings wird ein weiterer Jahrbuch-Schatz gehoben. Er ist durch seinen Bezug sowohl zum Freiheits- wie auch zum Gottesproblem von Bedeutung und wird von Matthias Lutz-Bachmann (Frankfurt/M.) kundig eingeführt und kontextualisiert. Mit seinen Repliken auf die kritischen Diskussionsbeiträge vom letzten Heft beschließt Luciano Floridi (Oxford) die sechste Jahrbuch-Kontroverse über „A new political ontology for a mature information society“.

Volker Gerhardt

# ‘Ultimate Responsibility’ without *causa sui*

Schelling’s Intelligible Deed of Freedom contra Galen Strawson’s Argument\*

Thomas BUCHHEIM (Munich)

*Abstract.* Since the mid-1980s, Galen Strawson has introduced an argument into the analytic debate about the concept and possibility of freedom. He has repeated and defended it in various formulations, which amounts to an “impossibilism” of freedom in the moral sense, i.e., to the impossibility that we can be called ultimately responsible for the moral quality of our actions based on existing freedom in the full sense. In this paper, I want to explain Strawson’s argument, which is supposed to prove this intuitive difficulty as *impossible* to fulfill, and to show the conditions of its persuasiveness. Furthermore, I will make clear how and by what right philosophers like Kant, Fichte and especially Schelling were able to evade this argument *avant la lettre* by introducing the concept of an intelligible self-constituting act of freedom.

Since the mid-1980s, Galen Strawson has introduced an argument into the analytic debate about the concept and possibility of freedom. He has repeated and defended it in various formulations, which amounts to an “impossibilism” of freedom in the moral sense, i.e., to the *impossibility* that we can be called responsible (‘truly’ or ‘ultimately responsible’) for the moral quality (right or wrong, good or evil) of our actions based on existing freedom in the full sense. As an example and proof of this, I would like to cite only one typical sentence of Galen Strawson from his paper “The Impossibility of Ultimate Moral Responsibility” summing this up:

It is exactly as just to punish or reward people for their actions as it is to punish or reward them for the (natural) colour of their hair or the (natural) shape of their faces.<sup>1</sup>

Strawson’s argument is eminently suitable, on the one hand, for pointing out an aporetic difficulty in our ordinary conceptions of human freedom and the moral demands we associate with it. On the other hand, it seems to be especially suited for gaining a comparative measure of how classical theories of freedom, particularly from the circle of Kantian and post-Kantian philosophy, have sought to resolve this difficulty before it was even condensed into a denial of our moral ultimate responsibility based on freedom by an argument like the one Strawson puts for-

---

\* I wish to thank Jörg Noller and Inken Titz for the translation of the article into English, done with great linguistic and philosophical expertise. I also thank the anonymous reviewers of this journal for critical comments and suggestions on an earlier version of the paper.

<sup>1</sup> Strawson (2008), 326.

ward. The intuitive difficulty mentioned above consists in the fact that we can be held (ultimately) responsible for an action only if its execution can be traced back to the agent as the decisive cause. Further, this being the decisive cause must not possibly be traced back to other sources than the ones from which also the action to be answered for originates. This requirement of a closed grounding of responsibility of morally relevant acts in the agent itself is not easy to fulfill for beings like us, who are born and will die.

First, I want to explain Strawson's argument, which is supposed to prove this intuitive difficulty as *impossible* to fulfill, and to show the conditions of its persuasiveness. Secondly, I will make clear how and by what right philosophers like Kant, Fichte and especially Schelling were able to evade this argument *avant la lettre* by introducing an intelligible self-constituting act of freedom.

### 1. Reconstruction of Strawson's argument

Strawson has presented the argument in many slightly modified versions, but always distinguishing a 'basic form' from a more elaborate 'cumbersome' form.<sup>2</sup> The 'basic argument' directly names the core point which is decided in it and which freedom in the moral sense, i. e., the reclamation of moral ultimate responsibility for one's own actions, cannot bypass in his opinion. This core point consists in what Strawson considers the indispensable requirement of being a *causa sui* as a bearer of ultimate responsibility for one's actions. But this, he argues, is impossible, at least for any finite being. I quote a short version of the 'Basic Argument' from "The Impossibility of Ultimate Moral Responsibility":

(1) Nothing can be *causa sui* – nothing can be the cause of itself. (2) In order to be truly morally responsible for one's actions one would have to be *causa sui*, at least in certain crucial mental respects. (3) Therefore nothing can be truly morally responsible.<sup>3</sup>

Immediately, of course, the attention will be directed to the second step of the argument with the question whether, even if only "certain crucial mental respects" of an individual claiming freedom for himself are concerned, the concept of a *causa sui* is to be applied at all and whether this does not rather mean to aim with cannons at sparrows? For example, we do not hesitate to attribute to the human being a radical acquisition of language ability or the self-transformation toward a good gymnast or strong wrestler without using a *causa sui* claim. So why have to be

<sup>2</sup> For a detailed presentation and contextualization of the argument, drawing on the objections and support it has received in the contemporary discussion of the concept of freedom, see Dettinger (2015), 66–112. However, Dettinger largely refrains from initially precisely highlighting and critically examining the argument's internal premises, on which the argument's (in any case informal) conclusiveness is based. Without this, however, it is not possible to identify sharply enough where alternative views on the matter would have to be hooked in order to deprive it of its persuasive power. Instead, Dettinger accepts it on the whole to make plausible an ultimately theological thesis about only "eschatologically" possible freedom of man.

<sup>3</sup> Strawson (2008), 319.

*causa sui* in the ‘certain respects’ of moral responsibility? In another essay<sup>4</sup> Strawson himself writes:

The claim, then, is not that people cannot change the way they are. They can, in certain respects [...]. The claim is only that people cannot be supposed to change themselves in such a way as to accede to UR with respect to the way they are, and hence with respect to their actions. One can put the point (somewhat contentiously) by saying that in the final analysis the way you are is, in every last detail, a matter of luck – good or bad.<sup>5</sup>

The source of the action for which someone has ultimate responsibility (UR) must not have been produced by a string of fortunate fates, but only on a track that in turn is at the responsible disposition of the agent. The crucial point in the case of moral freedom is thus, according to Strawson, that the self-change cannot lead to the *special* result of a person’s being *responsible* for the way she is. For always, according to the argument, what we do follows from the way we are, and consequently what we do *responsibly* follows from the way we are. But if what we do, *as a consequence* of the way we are, is to be something for which we are *responsible*, then we must *also* already be responsible for that from which it followed, i. e., for the way we are. So it seems to be intuitively quite plausible at any rate:

- (1) You do what you do because of the way you are.
- (2) To be truly morally responsible for what you do you must be truly responsible for the way you are – at least in certain crucial mental respects.<sup>6</sup>

It is important that we grasp more precisely the sense of the derivative or entailment relation between the being of the agent and the action in question, which Strawson’s argument must assume in all its versions in order to acquire its compelling force. He refers to this connection himself in various terms. Most often he speaks of “because of” or just “*causa sui*”, but it is clear that this does not mean causality in the ordinary sense. For in the usual understanding a ‘cause’ is first to be described logically independent of the caused effect and secondly takes place ahead of it. Neither is the case here: the constitution (“the way you are”) is one of the *same* subject that commits the action, and it is at the same time with it the cause from which it results.

Yet Strawson also uses different and even stronger expressions for the connection, such as “what one does is a function of how one is”<sup>7</sup> or “flows necessarily from how you are”<sup>8</sup>. However, it is not justified without further ado to recognize in this already a not only clearly functional but a necessary connection (“necessarily”). For, given the same external circumstances, it can probably not be called *impossible* that an agent somehow acts differently<sup>9</sup> than he does in the factually given case.<sup>10</sup>

<sup>4</sup> “The Unhelpfulness of Indeterminism” (Strawson 2000).

<sup>5</sup> Strawson (2000), 151.

<sup>6</sup> Strawson (2008), 325. The further steps are also quoted from this paper.

<sup>7</sup> Strawson (2008), 319.

<sup>8</sup> Strawson (2008), 325.

<sup>9</sup> It cannot be ruled out, for example, that in another possible world with the same external circumstances the person concerned would take more time to look at the photo more closely.

<sup>10</sup> Here it is neither implied nor excluded that the action someone actually does is subject to determinism or else indeterminism of its occurrence: For, according to the assumption I have made, it is only the

For example, someone *could* recognize his schoolmate on a photo shown to him, if he took more time to focus on the face on which he only casts an all too fleeting glance. Nevertheless, of course, de facto, i. e. in the real world, his statement that he did not know the depicted person resulted from all the individual traits of his constitution when looking at the picture. However, not taking more time for it does not *necessarily* (but only factually) belong to the feature of his engagement with the action. In a later version of his argument, Strawson elaborates on the entailment relation in question:

Consider a particular action or piece of deliberation in which you engage, and consider everything about the way you are when you engage in it that leads you to engage in it in the way you do. I will call the particular action or piece of deliberation that you engage in 'A', and I will call everything about the way you are mentally when you engage in it that leads you to engage in it in the way you do 'N'. [...]

3.1 When you act or deliberate, at  $t_1$  – when A occurs, at  $t_1$  – you do what you do, in the situation in which you find yourself, because of the way you are – because you are N, at  $t_1$ .<sup>11</sup>

From this, it becomes quite clear what we observed above that antecedent condition and result state belong to the same subject, i. e., it is an internal derivative relation between real states, which can be of neither a logical nor a causal nature. The former not, because real states do not enter into logical relations. The latter not, because both A, i. e. *actio*, and N, the ontological state from which the *actio* results, do not relate to each other in a time-differentiated way, but simultaneously: Strawson speaks of “the way you are [...], when you engage in”; thus, the so being is precisely when the *actio* is. Strawson seems to link to a scholastic *adagium* of the same content here, namely *actio sequitur esse* for any substance. The latter is an internally asymmetrical derivation relation. As such, it is clear from this that Strawson can assert a direct and unobstructed connection, but not logical or causal-nomological necessity for the entailment relation. For there would be no contradiction if this single subject did not engage in precisely the action in question, but in another one which, under slightly modified external circumstances, could be equally connectable to a given state of being of the agent. Accordingly, the entailment relation can be understood as well-founded emergence of the derived ‘engagement’ from the assumed antecedent, i. e. the being of the subject, which is here denoted by “N” as in “Nature”.

Second, it should be noted that Strawson in the quoted passage explicitly focuses on all “mental” aspects of the ontological constitution of one and the same subject, so that the exhaustive internal and therefore temporally and spatially unseparated ground of the action (or “piece of deliberation”) is placed in any *mental* states of being of the agent. Consequently, we have to understand the connection in question

---

‘external’ circumstances of the action that are assumed to be the same, while nothing is said about internal antecedents. Moreover, it became clear from the beginning that Strawson’s argument does not concern at all the question of the compatibility or incompatibility of freedom and responsibility with determinism (cf. also Strawson 2000, 151: “Note that the argument is completely independent of any view about whether determinism is true or false.”).

<sup>11</sup> Strawson (2002), 444 sq.

as a modern formulation of what one liked to call “psychological determinism” in classical philosophy.

Meanwhile, thirdly, what must certainly be called *subjectively* ‘impossible’ is that the same subject can escape an internal determination or dependency relation between two simultaneous but asymmetrically coupled aspects of itself. That is, it must be unavailable or *indispensable* for the subject itself that the *actio in* question follows its *esse*.

At this point, the transfer of responsibility from the action to the being of the person becomes plausible also in a logically more stringent, no longer purely intuitive sense. For *if* the subject is supposed to be responsible for her action, and the action at the same time results (in the way explained) from something that this subject cannot possibly prevent, then it follows logically, i.e. it is implied, that the subject must be responsible for that from which for her unescapably her action results. Therefore, we can accept *transfer* as a fixed building block of Strawson’s argument. This means that the subject could be responsible for his action *A* only *by being* responsible also for his mental state of being (*N*) at the time of the action.

Thus, by means of the transfer principle, we can now concede and reformulate the first two premises of Strawson’s argument, namely:

- |                                |  |
|--------------------------------|--|
| ( <i>actio sequitur esse</i> ) | (1) You do what you do because of the way you are. <sup>12</sup>   |
| ( <i>Transfer</i> )            | (2) To be truly morally responsible for what you do you must be truly responsible for the way you are – at least in certain crucial mental respects. <sup>13</sup> |

Now Strawson’s claim is that it is *impossible* for a subject (*S*) to be ultimately responsible (*UR*) for any state of being (*N*) at any time (*t<sub>n</sub>*). However, if this were impossible, then everything for which this is a necessary condition would be impossible. This is stated by proposition (2) of the argument.

In order to prove this impossibility, Strawson assumes the opposite in a follow-up argument, i.e. the *possibility* of being ultimately responsible for one’s relevant nature at any time, in order to show in the next steps that this assumption leads to an inevitable infinite regress. Since such a regress is again to be considered inadmissible according to Strawson,<sup>14</sup> but it must follow from the assumption (according to the argument put forward), it cannot be otherwise than that the assumption is to be rejected. This is what we want to follow with due brevity. Therefore, we set as target assertion:

- (3) You cannot be truly responsible for the way you are, so you cannot be truly responsible for what you do.<sup>15</sup>

<sup>12</sup> Strawson (2008), 325.

<sup>13</sup> Loc. cit.

<sup>14</sup> Such a regress is not per se logically inadmissible or contradictory. However, since the present case is about a foundational context for accountability, it can be said that an infinite regress is incapable of providing a completed foundational context (see, e.g., Schaffer 2010, e.g., 37; 62), and for that very reason the assumption that implies it must be rejected as unsatisfiable.

<sup>15</sup> Strawson (2008), 32.

It is important to note here that Strawson introduces another premise as a justification for impossibility, which apparently only establishes the meaning of what it *minimally* means to be responsible for something. Namely, it means to have intentionally produced that for which one is responsible:

(*N intentionally brought about*) Why can't you be truly responsible for the way you are?

Because (4) To be truly responsible for the way you are,  
you must have intentionally brought it about that you are  
the way you are, and this is impossible.<sup>16</sup>

What to make of this additional premise? It looks innocent at first sight, but it is not at all. For the intentional production of *N* must have two features according to the additional premise repeatedly mentioned by Strawson. It must both have come to completion in *N* and have taken place prior to *N*: *before* anyone can be responsible for his action, he would have to have intentionally brought forth from himself at an *earlier* time the nature *N* that now constitutes the foundational ground for his responsible action. We saw earlier that the internal and for *S* indispensable, simultaneously and inseparably occurring resulting of the *actio* from the *esse* of the person are only partial, but at the same time asymmetrically interdependent state aspects of an acting subject. According to the premise now introduced, however, they are mutually independently occurring total constitutions or real 'stages' of a person on an assumed path of her development through time. It is now *this* additional assumption that leads with rapid steps into regress. Namely as follows:

Why is it impossible? Well, suppose it is not. Suppose that

(5) You have somehow intentionally brought it about that you are the way you now are, and that you have brought this about in such a way that you can now be said to be truly responsible for being the way you are now.<sup>17</sup>

Now the regress begins, since having any intentions or preferences to do something intentionally, according to Strawson with proposition (1), always presupposes some *already given* being of the same subject, which has or pursues the intentionality in question. But since now the intentional bringing forth of the *N*-from-*S* must have come to a conclusion earlier than the bringing of this *N* into the performance of that action for which a responsibility is asserted, that which should only be given *by* the intentional bringing forth would have to be already presupposed for the intentional bringing forth. Thus proposition (6) of the argument results:

For this to be true

(6) You must already have had a certain nature *N* in the light of which you intentionally brought it about that you are as you now are.<sup>18</sup>

And here we get into the infinite regress. For always, in order to be responsible for any state of being *N*(*t<sub>n</sub>*), we have to assume, according to proposition (4), its intentionally having been brought about by the same subject (*S*), which possesses this

<sup>16</sup> Strawson (2008), 325.

<sup>17</sup> Loc. cit.

<sup>18</sup> Loc. cit.

nature.  $SN(t_n)$  is therefore in any case the *result of* an intentional bringing about of S. But in order to be allowed to assume just this intentional bringing about, we must (according to proposition (1): *actio sequitur esse*) at the same time *internally* presuppose that it already possesses a nature exhaustively founding the bringing about. Some *previous* nature  $N(t_{n-1})$  of S is therefore also a necessary precondition of the intentional production of  $SN(t_n)$ . However, again according to proposition (2) – i. e. by the *transfer* of responsibility – that S can be responsible for an intentional bringing forth only if it is in turn also responsible for the being to be presupposed internally for it or the nature  $N(t_{n-1})$  exhaustively founding the intentional bringing forth. The latter, however, requires again (according to proposition (4)) that *also* the nature  $N(t_{n-1})$ , *by virtue of* which the subject intentionally brought forth the nature  $N(t_n)$ , for which it must be responsible, if it is to be responsible for the action (A) resulting from it, must in turn have been intentionally brought forth by it. Etc. Thus, proven by the individual propositions of the argument, for every existing responsibility of the subject for an intentional action, there must always be switched on by it another nature already previously established by its intentional action, for which it must in turn draw responsible by its intentional action related to it.

This is how sentence (7) of the argument expresses it, namely:

But then

(7) For this to be true you and you alone are truly responsible for how you now are, you must be truly responsible for having had the nature *N* in the light of which you intentionally brought it about that you are the way you now are.

So

You must have intentionally brought it about that you had that nature *N*, in which case you must have existed already with a prior nature in the light of which you intentionally brought it about that you had the nature *N* in the light of which you intentionally brought it about that you are the way you now are ...

Here one is setting off on the regress. Nothing can be *causa sui* in the required way.<sup>19</sup>

It is, as one now clearly recognizes, not only the *transfer* of responsibility from the action to the being of the agent, which entails an infinite and insofar erroneous founding regress. Instead, it is the transfer together with the assertion that responsibility always means to have intentionally produced that being to which the transfer refers back, in order to be able to redeem the responsibility implied by the transfer. This, however, will turn out to be a not necessary additional assumption: the intentionality of an action, which we do responsibly, is not that, which, if necessary, would establish our moral ultimate responsibility for the ontological constitution, to which the declared transfer of responsibility leads back. So it will show our look at Schelling's doctrine of man's intelligible deed of self-determination in the *Freedom Essay*. For Schelling does admit the first two premises of Strawson's argument with certain modifications. But his conception of a transcendental constitutional act, by which every human being establishes their moral responsibility and freedom, does not result in an infinite regress, which would prove the impossibility of moral ultimate responsibility.

<sup>19</sup> Strawson (2008), 325 f.