Erkenn (2018) 83:61–87
https://doi.org/10.1007/s10670-016-9872-0

ORIGINAL RESEARCH

Freedom and the Phenomenology of Agency

Martine Nida-Rümelin[1]

Received: 18 December 2016 / Accepted: 18 December 2016 / Published online: 10 February 2017

Abstract Free action and microphysical determination are incompatible but this is so only in virtue of a genuine conflict between microphysical determination with *any* active behavior. I introduce active behavior as the veridicality condition of agentive experiences (of oneself as active) and of perceptual experiences (of others as active) and argue that these veridicality conditions (a) are fulfilled in many everyday cases of human and non-human behavior and that they (b) imply the incompatibility of active behavior with microphysical determination. The main purpose of the paper is to show that the view proposed about active behavior leads to a natural compromise between libertarianism and compatibilism, which avoids the flaws of both positions while preserving their central insights.

1 Introduction

In contemporary philosophy of agency it is considered a central task to account for the difference between actions and other events. But there is, according to the view here presented, a more fundamental distinction to be drawn: the one between active behavior (bodily or mental events brought about in a sense to be explained by the subject concerned) and mere happenings (bodily or mental changes the subject at issue passively undergoes). According to the view here proposed, the capacity to behave in an active manner is not restricted to the human domain and might already have occurred with the emergence of simple forms of consciousness in early stages of biological evolution.[1]

[1] A similar view is proposed by Steward (2012a, b). However, contrary to the position here proposed she claims that agency in the relevant broad sense here called 'active behavior' requires a 'two-way-power' in a sense which implies the metaphysical possibility of more than one course of action.

✉ Martine Nida-Rümelin
martine.nida-ruemelin@unifr.ch

[1] Université de Fribourg, Fribourg, Switzerland

It is hard to deny that we all share a vivid understanding of the distinction between changes that happen to a subject and changes the subject actively brings about. But how do we form this understanding and what is it based upon? According to the view here proposed our main cognitive route to the difference between activity and passivity is phenomenological: *we understand what it is to be active on the basis of our awareness of being active while being active in our own doings*.

In daily life we experience our own doings as actively brought about by ourselves; they seem to be generated in a way which involves that *we*, the experiencing being we are, *actively* contribute to their occurrence. Analogously, we experience the behavior of others (of human subjects and of non-human animals we take to be conscious) as actively brought about *by them*. I will briefly present an argument for the following conclusions:

(C1) Experiences of our own present behavior as active and perceptions of the behavior of other conscious subjects as active are (in large part) veridical. In other words: conscious subjects *are* active in their behavior in the way in which they *appear* to be active in those experiences

(C2) A behavior which is active in the relevant sense (brought about by the subject concerned in the sense at issue) is not causally determined by previous events

The argument for these conclusions will be sketchy, it is in need of further elaboration and it is not the main theme of the present paper. I will focus here on a different issue: I would like to explore how these conclusions, if true, can help us to get clear about the old puzzle about freedom and determination.

One central ingredient of the view which will emerge is this: the much-discussed apparent or genuine tension between free agency and determination has been miss-localized. The supposed incompatibility of free action and determination has nothing in particular to do with freedom; it is rather due to the fact that *all* action (be it free or unfree) is *active* behavior.

Active behavior, as will be argued below, is *incompatible* with microphysical determination but it is *compatible* with determination in the sense of there being no *metaphysically possible* alternative to the way the agent acts in a given case. This distinction will prove helpful for developing a better understanding of the various competing core intuitions underlying the debate about freedom and determination.

2 Agentive and Perceptual Experiences of Behavior as Active

The following description cited from Horgan et al. (2003) is an excellent way, I find, to 'point' to that specific aspect:

> "…What is behaving like phenomenologically, in cases where you experience your own behavior as action? Suppose that you deliberately perform an action—say, holding up your right hand and closing your fingers into a fist. As you focus on the phenomenology of this item of behavior, what is your experience like? To begin with, there is of course the purely behavioral aspect of the phenomenology—the what-it's-like of being visually and

kinesthetically presented with one's own right hand rising and its fingers moving into clenched position. But there is more to it than that, of course, because you are experiencing this bodily motion as your own action. In order to help bring into focus this specifically actional phenomenological dimension of the experience, it will be helpful to approach it a negative/contrastive way, via some observations about what the experience is not like. For example, it is certainly not like this: first experiencing an occurrent wish for your right hand to rise and your fingers to move into clenched position, and then passively experiencing your hand and fingers moving in just that way. Such phenomenal character might be called the phenomenology of fortuitously appropriate bodily motion. It would be very strange indeed, and very alien. Nor is the actional phenomenological character of the experience like this: first experiencing an occurrent wish for your right hand to rise and your fingers to move into clenched position, and then passively experiencing a causal process consisting of this wish's causing your hand to rise and your fingers to move into clenched position. Such phenomenal character might be called the passive phenomenology of psychological state-causation of bodily motion. People often do passively experience causal processes as causal processes, of course: the collision of a moving billiard ball with a motionless billiard ball is experienced as causing the latter ball's subsequent motion; the impact of the leading edge of an avalanche with a tree in its path is experienced as causing the tree to become uprooted; and so on. But it seems patently clear that one does not normally experience one's own actions in that way—as passively noticed, or passively introspected, causal processes consisting in the causal generation of bodily motion by occurrent mental states. That too would be a strange and alienating sort of experience.

How, then, should one characterize the actional phenomenal dimension of the act of raising one's hand and clenching one's fingers, given that it is not the phenomenology of fortuitously appropriate bodily motion and it also is not the passive phenomenology of psychological event-causation of bodily motion? Well, it is the what-it's-like of self as source of the motion. You experience your arm, hand, and fingers as being moved by you yourself—rather than experiencing their motion either as fortuitously moving just as you want them to move, or passively experiencing them as being caused by your own mental states. You experience the bodily motion as caused by yourself."

In the cited passage a preliminary account of how things appear in agentive experiences with the relevant aspect is proposed: the agent is under the impression of causing him- or herself the relevant physical event (the movement of the hand). At the same time the passage has a different function: it aims at attracting attention to a specific feature or aspect of agentive experience. It thereby tries to establish shared reference among the author and the reader to a particular phenomenal aspect of normal human agentive experience. It is the second function of the text which is most relevant for my present purposes. Despite a number of objections one might raise against the contrasting characterization and against the particular formulation of the supposed content of the experience, the passage is admirable, I find, in the

way it attracts attention to a specific aspect of agentive experience.[2] For most readers, I believe, it is successful in establishing shared reference to that omnipresent aspect of the way it is like to act. In the following I will assume that this is so.

What is called in the cited passage "the experience of self as source" is precisely what I mean saying that we experience our own behavior *as actively brought about by ourselves*. I prefer the latter terminology for the following reason: agentive experience with respect to one's own behavior and perceptual experience concerning the behavior of others have the very same kind of content. Therefore, we need to introduce a terminology that allows us to describe that common content in a natural manner. Describing that content saying that some piece of behavior appears to be actively brought about by the subject concerned (ourselves in the agentive case, the perceived other subject in the perceptual case) allows us to do so. By contrast, to say that we perceive "other selves as source" would be a quite unhappy way to put the point since it would misleadingly suggest that we see other selves. Furthermore, it seems quite important to me to realize that there is a close relation here between the way our own behavior is given to us (in acting or in doing something) and the way the behavior of others is given to us (in observing them as they do things) which should be kept in mind and underlined already by the terminology one chooses.

Here are a few more examples of *agentive* experiences with the aspect of 'experiencing one's own behavior as active'. Movements done without any conscious intention and in an inattentive manner like moving one's legs while listening to a talk are experienced, nonetheless, as brought about by ourselves. Complex movements such as dancing, even if done in a totally unplanned manner and in a way one experiences as not consciously controlled, are experienced as actively brought about by oneself. Some elements of our mental life are experienced as actively brought about by ourselves. This is typically so for conscious decisions taken on the basis of deliberation (by contrast, mental events such as thought contents suddenly popping up in one's mind are not experienced as actively brought about by oneself). Quick reactions without reflection to a suddenly perceived danger (for instance while driving a car) are experienced as actively brought about even though they are not experienced as done for reasons that are present to the person in the moment of the reaction and even though they are experienced as 'directly triggered' in some sense by the perception of danger.

Here are a few examples of *perceptual* experiences of active behavior in others: when a person you talk to suddenly smiles in an amused manner then you see the smile as an active behavior of that person, the person you are confronted with seems to be actively involved in the way that smile comes about. The smile does not look to you like the result of microphysically determined inner neural and muscular processes. It looks as if done by the person in a way which excludes such microphysical determination, or so I would like to suggest. Similar observations apply to seeing animals in their behavior. You see the bird as if actively moving its

[2] See for critical remarks on the description given here chapter 8 in my forthcoming book "Experiencing subjects. Sketch of a theory".

head when it turns its gaze towards you and you see the squirrel, the experiencing subject you are confronted with on such an occasion, as running up the tree in an active manner, where that active manner involves, according to the way things appear to you, that it is the squirrel itself who actively contributes to its running.

3 Veridicality Conditions of Experiences: Terminology and Presuppositions

Before starting to present a few more specific claims about the phenomenology of agency and the phenomenology of perception it will be helpful to introduce some terminology and to make several theoretical presuppositions explicit. Experiences as the term will be understood here are instantiations of *experiential properties* by experiencing subjects.[3] Experiential properties can be characterized as follows:

Definition 1 A property P is an experiential property iff (a) it can only be instantiated by experiencing subjects and (b) its instantiation by a subject S *partially consists in* what it is like for S to have property P.

Experiential properties in the sense of that definition need not be exhausted by what it is like to have them. For instance, having a headache is an experiential property which is not exhausted by what it is like to have it since having a headache involves having a feeling with a particular kind of cause, or so at least one may plausibly argue. Experiential properties in a strict sense (as opposed to the broad sense just defined above) are experiential properties such that having them is exhausted by what it is like to have them. To be acting is an experiential property in the broad but not in the strict sense. It is something like for the subject who acts to be acting but the fact that the subject is acting does not consist in what it is like for the subject to act.

It is characteristic of many experiences that the subject at issue is under the impression that certain things are the case. For instance, in being visually presented with a tree the subject is under the impression of there being a tree with specific properties in front. If there is in fact no tree in the subject's environment, then the experience is not veridical: it does not present the world as it actually is. Veridicality of an experience may be defined as follows:

Definition 2 (*Veridicality*) An experience of type E is veridical iff the way the world appears to be to the subject in having the associated experiential properties is the way the world actually is.

The above definition presupposes that for every experience type E there are associated experiential properties such that each occurrence of an experience of type E consists in the instantiation of those experiential properties by some experiencing

[3] The claim that issues about phenomenal consciousness should be approached in terms of experiential properties and not, as it is usually done, in terms of supposed qualitative properties of events called 'experiences' is elaborated and defended in Nida-Rümelin (2016a).

subject. An experience is veridical just in case its associated veridicality conditions (in the sense of the definition below) are fulfilled.

Definition 3 (*Veridicality conditions*) The conditions C constitute the veridicality conditions of a given experience type E iff in a case where an experience of that type E is veridical this is so in virtue of the satisfaction of those conditions C.

I will be assuming in what follows that phenomenal kinds of experiences can be described by the veridicality conditions that are common to all members of that type. I do not thereby subscribe to the stronger claim that phenomenal properties can be reduced to intentional properties. This stronger claim can be formulated within the terminology chosen here in the following way: for every experiential property P there is a specification of veridicality conditions C_P such that having property P consists in being under the impression that the conditions C_P are fulfilled. Even if one explicitly rejects the possibility of such a reduction one may still accept that one can at least partially describe in many cases what it is like to have a certain experiential property by describing what appears to be the case to someone who has that property. Since (a) phenomenal kinds of experiences are individuated by the associated experiential properties and (b) experiential properties can be described (even though not fully captured) by what appears to be the case to someone having the property and (c) the veridicality conditions of an experience specify what appears to be the case for someone having the associated experiential properties, it follows that one can describe phenomenal kinds of experiences by reference to the associated veridicality conditions.

For instance, one can describe the specific aspect of the experiential property of acting which has been called here "the experience of one's own behavior as active" by specifying how things appear to be to the agent in acting (and in *thereby* undergoing an experience with that aspect). It seems phenomenologically obvious that these veridicality conditions concern *the way the behavior comes about*; they concern the way the behavior is generated. The same applies quite obviously to the perceptual case. When you see the squirrel as actively contributing to its bodily movements then those movements appear to be generated in a specific way. To specify the veridicality conditions of these experiences is to specify what appears to be the case to the subject undergoing such experiences. I will assume here and I hope that this will seem obvious to most readers: the conditions that must be satisfied by a given behavior for the experience of that behavior as active to be veridical are the same for experiencing one's own behavior as active (in the agentive case) and for experiencing the behavior of others as active (in the perceptual case).

4 The Veridicality Conditions of Experiences of Behavior as Active and the Nature of Being Active

I would like to summarize the view her to be explored by the slogan that *conscious animals are genuinely active in their behavior* and I will call it the GAA view accordingly (derived from "**g**enuinely **a**ctive **a**nimals"). To add "genuinely" in its

characterization is meant to point to the fact that being active is understood here in a specific sense. I *don't* mean to call a behavior active for the simple reason that internal states of the animal contribute substantially to its occurrence. I mean to call it active only if there is a sense in which the animal as such brings the behavior about and this *precisely in the way which renders our everyday experiences of our own behavior as active and the behavior of other animals as active veridical*.

But why should one assume that to be active *is* to fulfill the veridicality conditions of the relevant experiences of being active? Here is the motivation: our primary cognitive access to what it *is* to be active is experiential: we understand what it is to be active via our agentive and perceptual *experiences* of being active. If those experiences are all illusionary, then there is no such thing as active behavior. It is therefore justified to introduce the phenomenon of being active (the phenomenon referred to by "active behavior" if there is such a phenomenon) *by stipulation* as the condition which renders experiences of being active veridical. Accepting this stipulation implies that the way we *experience* items of behavior to be generated when we experience them as active puts constraints on the very nature of the phenomenon we talk about using the term "active behavior". If, for instance, to experience an item of behavior as active involves experiencing it as *not* caused by previous events, then active behavior (if it exists) is *not* caused by previous events.

To determine the correct description of the veridicality conditions of a given type of experience is not a trivial task at all nor is it obvious how one should proceed to find an adequate description in a given case.[4] There is no room here to weigh different possible proposals against each other for the case here at hand of agentive and perceptual experiences of behavior as active. I will simply work with the following suggestion and leave defending it against objections to a different occasion:

(PH1) In experiencing ourselves as active in behavior B, we experience the behavior as brought about by ourselves in a way which excludes that the behavior is causally determined by previous events (such as occurrent desires or wishes or by the brain processes 'realizing' them).

(PH2) In experiencing the behavior of other animals as active, we experience their behavior as brought about by them in a way which excludes that their behavior is causally determined by previous events.

(PH1) should not be misunderstood as claiming that our own mental states and other events such as physical processes in our brain appear to play no role in the generation of our own behavior. To the contrary one should of course admit that in many cases of agentive experience we are not only aware in a phenomenally manifest manner of actively contributing to the realization of certain movements but rather also aware in a phenomenally manifest manner of what *motivates* us to actively do what we do. It follows from (PH1) that the relevant experiences are veridical only if motivation is distinct from event causation. In other words: for a

[4] The methodological issue about how one may determine the content of experiences which does duty to phenomenology while at the same time taking into consideration theoretical constraints is discussed in Siegel (2012), chapter 3 and in Horgan (2007).

conscious person to be motivated by reasons must not be equated with the relevant item of behavior being causally determined by events such as realizers of the relevant representation of reasons in the brain. According to (PH1) the agentive experience of being active is illusionary if for an item of human behavior to be motivated by certain considerations *is* for that item to be caused by appropriate brain processes some of which are the physiological basis of the relevant conscious considerations. Rather, if these experiences are veridical, then something like this must be an adequate description of rational human behavior: the person at issue *actively* brings about behavior B *in light* of certain considerations. Even if these considerations are so compelling that the person has *no other choice* given her values and convictions the behavior the person brings about *actively* is not causally determined be previous events.

Similar remarks apply to the content of perceptual experiences of behavior as active. (PH2) should not be misunderstood as claiming that seeing the squirrel as actively involved in its running includes the impression that the squirrel's fear does not play any role. To the contrary, it may on a given occasion be part of the content of such a perceptual experience that you see the squirrel as actively bringing about its hectic running *because* it is frightened by your presence. (PH2) only implies the following: if your perception of the squirrel is veridical, then, yes, it runs *because* it is frightened, but the "because" here at issue must not be understood as stating a simple event-causal link such as the fear or its basis in the brain mechanically causing (together with other brain processes) the running movements. To say it in a vivid manner: according to (PH2) you see the squirrel's movement in a way which excludes that the squirrel is a biological mechanism. According to the way you see it, the squirrel is a conscious active individual in a sense which excludes that it is a biological robot, a conglomeration of cells the behavior of which is microphysically determined.

The GAA view assumes that our experiences of behavior as active *are* in the overwhelming majority of paradigmatic cases veridical. It states, in other words, that in the large majority of paradigmatic cases the veridicality conditions associated to experiences of behavior as active are satisfied. In order to make the view explicit one has to characterize the veridicality conditions at issue. In line with what has been said so far, I propose the following claim about the veridicality conditions for experiences of behavior as active (it makes no difference if the experience is agentive or perceptual):

V: An item of behavior B fulfills the veridicality conditions of experiences of behavior as active if and only if *B is brought about by the subject concerned in such a way that B is not causally determined by previous events*.

It seems safe to assume that a bodily movement which *is* causally determined by previous events is microphysically determined. It therefore follows from V that no microphysically determined behavior fulfills the veridicality conditions of experiences of behavior as active.

The GAA view may now be characterized as follows: in the vast majority of paradigmatic cases of human and non-human behavior which are experienced as active the right hand side of the bi-conditional formulated in V is satisfied. The behavior is generated in the way it seems to be generated: it is actively brought

about by the subject in a way which excludes that it is causally determined by previous events and in a way which excludes that it is microphysically determined.

The above characterization is only a minimal description of the GAA view about which a lot more could be said. For the purposes of the present paper it is crucial to see that the GAA view allows for and even requires distinguishing two senses of determination: causal determination of concrete occurrences of items of behavior by previous events on the one hand and metaphysical determination of the *kind* of behavior realized at a given occasion on the other. The way a person acts may be metaphysically determined in the sense that there is no metaphysically possible counterfactual situation in which the same relevant preconditions are fulfilled and yet the person acts otherwise. And yet, according to the GAA view, even if a person's way of acting is *metaphysically* determined the person's act is not *causally* determined by previous events. Such situations of metaphysical determination can be adequately described in retrospect as follows: it was *within the causal powers* of the person to act otherwise but there was *no metaphysical possibility* that she would do so. This claim may appear paradoxical at first sight but it will, I hope, gain some plausibility for the reader when illustrated below by several cases. The distinction between causal and metaphysical determination here at issue will be crucial for the resolution of the puzzle about freedom and determination proposed in the last five sections of this paper.

Why should we accept the GAA view? Here is the answer I would like to propose: if what has been said about the veridicality conditions of experiences of behavior as active is correct, then the only alternative to the GAA view is an illusion theory. Omnipresent experiences concerning our own behavior and the behavior of others would have to be illusionary. When, for instance, you see the smile of a person as brought about by her, then this is (if the GAA must be rejected) strictly speaking a mere illusion. If, under the influence of some bad drug, you were to see the smile as a mechanistically caused facial change then, if the GAA is false, your experience of the smile would be veridical. If you see a squirrel as a conscious individual bringing about its running in an active manner, then, strictly speaking, you are, again, a victim of illusion. In reality, the squirrel is (if the GAA is false) a microphysically determined system; it is what one may properly call a biological robot. According to such an illusion theory the experience of being actively involved is an illusion created in us by certain brain mechanisms and the perceptual experience of others as active is a projection of that illusion onto others we are presented with in perceptual experience.

Is the illusion theory credible at all? That much should be clear: if there was a person who perceived others in the way which *would* be veridical according to the illusion theory, then he or she would live in a regrettable situation. One would consider his or her experience as pathological and would have deep pity given the sad impoverishment of his or her subjective world. This does not show that the illusion theory is false but it does show that accepting it is a high price to pay. We should be ready to pay such a price only if compelling reasons can be given to the effect that the GAA view is unacceptable. Those reasons can only be scientific reasons. Are there such reasons? I am convinced there are not, but this is a big issue which cannot be discussed here.

The argument sketched here for the GAA view can be attacked in a number of ways:

1. One may argue that there are no experiences of being active or that their veridicality conditions have been miss-described.
2. One may argue that there are compelling scientific reasons for accepting that those veridicality conditions are never fulfilled and go for the illusion theory.
3. One may argue that the illusion attributed to us by the illusion theory is not as radical and fundamental as has been suggested by the above description and that it is not so hard to accept that we are, in that respect, permanently the victim of illusion.

I don't believe that any of these objections is in the end successful. But I cannot even start to answer them here. Instead, I will explore the view about freedom and determination which emerges if we assume for the moment, for the sake of argument, that the GAA view is correct.

Before doing so it will be helpful to put the GAA view in the context of the contemporary landscape of available theories about agency. The reader will have recognized elements of what has been called the thesis of agent causation and some will wonder about the relation between such much-discussed views and the one here proposed. So let me briefly clarify this relation before I finally start talking about freedom.

5 The GAA View and Agent Causation

The GAA view is a close relative of the *agent causation view* (called the ACV in what follows). According to the ACV there is a specific kind of causation at work in free human action. While normal causation relates events as causes and effects, that special kind of causation relates individuals (a person, an agent) to events (e.g. bodily movements or changes in the brain). The person itself is considered to be the cause of the event at issue.[5] The ACV view is not standardly motivated by any phenomenological observations about how things appear to be in agentive or perceptual experience; it is standardly motivated by theoretical considerations about moral responsibility and human freedom. Yet, when pressed to say what agent causation is supposed to be some authors claim that we are familiar with it on the basis of our own action.[6] If one is attracted by agent causation theories at all, then one will naturally assume that it is the first person experience of bringing one's own movements about which gives us access to what it is for an agent to cause an event in the relevant sense of agent causation. This is, I take it, what makes the two views, the GAA view and the ACV close relatives. They both owe their intuitive attractiveness in great part to the way things appear to be in agentive experience.

[5] The agent causation view is famously defended in Chisholm (1964). More recent work defending the view has been developed by O'Connor [see, for instance, O'Connor (1996)].

[6] Chisholm (1964) Sect. 9 may be read along these lines.

However, in order to motivate the ACV by the assumption that agentive experiences of being active are veridical one would have to make a strong claim about their veridicality conditions. Instead of the claim V proposed in the preceding section one would have to say this:

V': The veridicality conditions of agentive experiences of behavior as active are fulfilled if and only if the behavior is caused by the subject (in the specific metaphysical sense stated by agent causation theories).

V' implies V but an argument would be required to show that V implies V'. The GAA view only incorporates V and is therefore weaker and thereby less problematic. It is open for the possibility that ultimately some version of the agent causation view must be accepted. But there are other possible accounts (one of them briefly sketched below) of what it is to be active in one's behavior compatible with V that are *not* committed to the metaphysical thesis that subjects (agents) are causes which might in the end prove superior for theoretical and empirical reasons.[7]

Related to the difference just explained the following advantages of the GAA view compared to the ACV view merit to be pointed out:

1. It has been argued that the ACV is inacceptable for the following reason: contrary to events individuals are not dated. But how can an event which is dated (it happens at a certain moment) have a cause which is not dated (it is not, in the relevant sense, located in time)?[8] The GAA view as characterized so far may be elaborated saying that for a behavior to be active is for it to fulfill the following condition: the subject at issue is actively involved in the generation of the behavior. If, for instance, the behavior takes place in the time interval (m1, m2) then for that behavior to be active involves that the subject at issue is *active in that behavior* in a time interval including (m1, m2). No specific metaphysical claim about what it is to be active in a given temporally extended behavior is implied by that description (in particular, as already stated before, it is not implied that the subject is, literarily, the cause of that behavior).
2. The ACV is often presented in a way which invites the idea that the agent intervenes in a punctual or instantaneous manner, thereby triggering one thing to happen instead of another. The picture invited is that, when lifting my arm to vote, for instance, I cause some initial brain event to happen which then further develops all by itself, so to speak, and causes my arm to move. This picture is, however, not the one invited by phenomenology as will be clear by some reflection and it is dubious for independent reasons.—The picture invited by the GAA view is quite different. In agentive experiences we appear to be actively involved in the generation of the *whole* movement while it is executed. We do not seem to instantaneously trigger the movement at its very beginning. Therefore, since being active is, by stipulation, what renders these experiences

[7] In earlier publications (Nida-Rümelin 2007, 2016b) I claimed that the behavior being caused by the subject (subject causation) constitutes the veridicality conditions of the experiences at issue. I now think that this was too quick. I owe this insight to discussions with Emmanuel Baierlé, Franziska Müller, Jacob Naito and Julien Bugnon.

[8] The objection is due to Broad (1953). For comments on the objection see Horgan (2007, p. 326) and Keil (2014, p. 85).

veridical, and the GAA view states that they *are* veridical, actively moving one's arm involves that the subject is active in the generation of the physical event during the whole period of time in which it happens.

Besides these two differences it will be clear by now that there are further points where the GAA view and the ACV diverge:

3. While the ACV restricts the phenomenon of *agent causation* to the human realm or even to the special case of free human action the GAA view claims that *active behavior* is widespread in the animal kingdom and present as well in human and non-human behavior that one normally does not classify as action.
4. The ACV view is standardly motivated by theoretical considerations about freedom and responsibility. By contrast, the GAA view is motivated by phenomenological observations about how we experience ourselves in our active behavior and about how we perceive others in theirs.

6 Resolution of the Problem About Freedom: Description of Its Central Elements

From the perspective of the GAA view the famous problem about freedom and determinism presents itself in a new manner. It turns out that puzzling and famous problems dissolve. Furthermore, using the GAA view one can isolate, for two standard responses, for compatibilism and for libertarianism, what it is about them that makes them intuitively appealing and what their insights are. Within the GAA view, an account of human freedom can be developed which unites those insights. It will be helpful to start with an overview of how the GAA view manages doing that.[9]

According to the compatibilist there is no genuine conflict between the thesis of human freedom and the claim that all human behavior is microphysically determined. This claim is unacceptable for a proponent of the GAA view according to which *all* active behavior is generated in a way which is incompatible with micro-physical determinism. So with respect to this element the proponent of the GAA view sides with the libertarian. However, the motivation given within the GAA view for denying the compatibility of free action and micro-physical determinism is different. The libertarian motivates the relevant incompatibility claim by some strong version of the so-called PAP principle (the principle of alternative possibilities) according to which free action requires that relevantly different possibilities are open and up to the agent in moments of free choice so that he or she *could have acted otherwise* under the very *same* external and internal conditions. Contrary to this the proponent of the GAA view rejects the PAP principle and

[9] For a detailed presentation and elaboration of the various responses to the problem about freedom and determination see Fischer et al. (2007) and O'Connor (2002/2010). In the German-speaking discussion the compatibilist position is defended among others by Beckermann (2006), Spohn (2012 and Pauen (2004) while different versions of a libertarian view are defended by Keil (2014) and Nida-Rümelin (2005).

argues that free action is incompatible with microphysical determinism not in virtue of being *free* but simply in virtue of being an instance of *active* behavior.

The proponent of the GAA view can and should deny the version of the PAP principle endorsed by libertarians and concede to the compatibilist that there are many cases of genuinely free action where the agent could *not* have acted otherwise and that free action is quite typically *determined* by the reasons available to the agent in the moment of choice in the following sense: given those reasons and all other relevant conditions satisfied at the moment of choice there was no psychological and no metaphysical possibility that the agent would act differently. To admit this is compatible—within the GAA view—with the denial of microphysical determination of the relevant piece of behavior. The GAA theorist can distinguish two relevant senses of determination of an instance of human behavior (psychological or even metaphysical determination on the one hand and microphysical determination on the other) and use this distinction to integrate the following claims: (1) in many cases of free action the relevant behavior is psychologically or metaphysically determined. (2) In no case of (free) action (and in fact, in no case of active behavior), the relevant item of behavior is microphysically determined. By integrating (1) the GAA view can preserve an important part of the compatibilist intuition; by integrating (2) the GAA view can preserve an important part of the libertarian intuition.—Let us now have a closer look at how the GAA view can integrate the compatibilist core intuition without thereby abandoning the libertarian core intuition.[10]

7 Psychological Determination

The core idea of the libertarian view may be formulated as follows: in cases of free action the person determines which of various future courses of events will be realized. In such cases the following condition holds: there is a moment m before the choice such that there are nomologically possible future continuations of the actual course of events until moment m in which the person acts differently from the way she actually acts and that are yet exactly like the real course of events with respect to all relevant conditions obtaining before and at moment m. For instance, when a person raises her arm in order to vote for a candidate and thereby acts freely, then—according to the libertarian core idea—it was still nomologically possible (that is compatible with the laws of nature) at some moment m briefly before she took her decision that she would not raise her arm; it was possible that she acts otherwise (and thereby moves otherwise) under the very same internal and external conditions obtaining at moment m. More specifically, the libertarian does not want the agent's decision to be determined by the prevailing psychological factors and wishes to defend the intuition that in the relevant cases of free choice the person can decide in various different ways under exactly the same psychological conditions

[10] I would like to thank my partner Max Drömmer and our son Korbinian Nida-Rümelin for extended discussions about the topic of this talk during a stay in Sardinia in September 2015 which helped me a lot to elaborate the details of the view here presented.

(same beliefs, same values, same character etc). The libertarian intuition thus includes that free actions are never psychologically determined: if an act is free, then the person could have acted otherwise under the very same psychological preconditions. This is where the libertarian goes wrong. Many actions (or choices) that should be considered as free are psychologically determined. Nonetheless a core intuition behind the libertarian view can be preserved and explicated within the GAA view.[11]

The notion of psychological determination will play a central role in the following discussion. In order to contrast the GAA view with the libertarian view it will be helpful to use a strong notion of psychological determination of a person's action which can be explained, roughly, in the following way: under the same psychological conditions it was *metaphysically* impossible that the person acts otherwise. To be more precise, the relevant notion of psychological determination must be relativized to a moment m in time along the lines of the following definition:

Definition (*Psychological determination at a given moment m*) S realizes action A at moment m'; m is prior to m'.

B is psychologically determined at m (or: determined by the psychological conditions obtaining at m) iff:

S realizes action A at moment m' in all metaphysically possible counterfactual situations in which

(a) S is at m' in the same objective decision situation as in the real world (with respect to all conditions relevant for what S does at m').
(b) S has, at m', the same psychological properties as far as they are relevant to what S does at m'.
(c) All other conditions obtaining between m and m' that are relevant to what S does at m' are the same as in the real world insofar as they are independent of S's self-generated psychological development between m and m'.

Condition (c) excludes cases in which an external event such as a stone falling on the person's head prevents her from realizing A at m'. If such possible situations where quantified over in the above definition then an action realized in the real world would count as *not* psychologically determined if it is metaphysically possible that an external event of the kind just mentioned prevents the person from acting in the way she would have acted had the event not occurred. It is however necessary to restrict the conditions mentioned in (c) to those that are independent of S's self-generated psychological development between m and m' to avoid another problem. An act is *not* psychologically determined at moment m if (for instance) the agent's active direction of attention makes a difference for what he or she does at m' and if it is not metaphysically excluded at m that she directs her attention in this specific way (the GAA view allows for that possibility). Such a case which

[11] I oscillate between talking of free choices and of free actions in what follows but this does not harm in the present context.

otherwise would be a counterexample to the definition proposed is excluded by condition (c).[12]

With these clarifications we can now turn to the way in which the GAA view can incorporate the compatibilist core intuition. The basic idea will be this: while microphysical determination is not compatible with free action since it is not even compatible with being active, an action can be free and yet psychologically determined in the sense just defined.

8 Obvious Choices

The libertarian view has a serious problem with situations of obvious choices. It is characteristic of such situations that it is immediately obvious to the agent what it is he or she should do and that the agent has no reason and no motivation not to do what appears to be the thing to do in the situation at hand. Here are four examples of that kind:

Example 1 While Anton is sitting on a river beach he observes a 2 years old boy who is playing next to the river. The parents' attention is caught by a conversation with friends. There are no other people on the beach. Suddenly the boy stumbles and falls into the water. Anton immediately understands the situation. He jumps up, hurries to the river and saves the little boy grasping his arm and pulling him out.

Example 2 Beatrice is a teacher of mathematics. Every Monday morning Beatrice stands up at 8.00 and leaves her home to go to the school in her neighborhood in order to start teaching at 9.00. This is what she does on Monday, April 11, 2016. She doesn't think about whether or not she will go to school today. It is obvious to her as she wakes up that this is the thing to do. She realizes her plan without any deliberation.

Example 3 Christopher must undergo a surgery. He must choose between two possible ways for the medical doctors to proceed. After having done what he could do to gain all relevant information within reasonable limits of time, Christopher is convinced that the first procedure is more expensive, more painful and more risky than the second. There are no other relevant differences as far as Christopher knows between the two procedures. Christopher takes the obviously rational decision in favor of the first option.

Example 4 Daniela is part of a political committee which must decide by majority vote who of two candidates will be assigned a responsible position. Candidate A is, in Daniela's view, a reasonable person who is well suited for the job. Candidate B is a person with sympathies for neo-Nazi groups and is likely as far as Daniela knows to use the power associated with the position in a morally problematic manner. It is obvious for Daniela that she should vote for candidate A and so in the relevant moment she does so by raising her arm.

[12] For more detailed explications of condition (c) in the above definition compare Nida-Rümelin (2016b), Sect. 7.

With plausible additional assumptions all four cases are such that a third person who knows the agent and the situation well enough could predict with certainty (setting aside intervening factors that prevent the person from acting or that change the situation in some relevant manner) what the agent will do. Furthermore plausible additional assumptions make it even *metaphysically* impossible that the person acts otherwise than she actually does. Let us assume that Anton has a realistic perception of the danger for the boy and of the likelihood that the boy will die if he doesn't intervene and let us assume furthermore that Anton is a normal rational person. Under these psychological preconditions together with the assumption that nothing prevents Anton from acting, it is simply impossible that Anton does not act in order to save the child. The psychological preconditions do not determine how exactly he executes the action; they do not determine in what way he jumps up, with how many steps he runs to the river and which of the boy's arms he grasps in order to pull him out. But it is impossible that he acts in a way which does not fall under the type "trying to save the boy". In every counterfactual course of events in which Anton does not try to save the child in the same objective situation, either some event occurs before he can do what he intends to do (in contradiction to condition c in the definition of psychological determination) or Anton has suddenly become crazy, or he has very different values, or he is suddenly befallen of an irrational fear of drowning himself etc. Under the conditions at issue, a person who is psychologically like Anton in all relevant aspects will try to save the child. It is *metaphysically* impossible (not compatible with the *nature* of Anton's psychological states) that a person psychologically like Anton does not in that situation act like Anton.[13]

Analogous observations apply to Beatrice. Let us assume that Beatrice is a teacher who likes her profession and feels responsible for her pupils. Her lesson for that particular day is already well prepared. She is looking forward to her work with the group of young people. No unusual event or new information on that particular morning gives her any reason to depart from the original plan. Once again it seems clear: in that particular situation and if nothing unusual happens before the plan is realized, a person who is psychologically just like Beatrice in all mentioned respects will of course go to school and teach. Again, it is not determined by those preconditions in what way Beatrice realizes the plan. But it is determined by those preconditions that she does something which falls under "going to school in order to teach mathematics". Under those psychological conditions (and the other two conditions mentioned in the definition of psychological determination) it is metaphysically impossible that Beatrice acts otherwise because acting otherwise is incompatible with the very nature of the psychological properties we have been assuming Beatrice to instantiate at the relevant moment in time.

Likewise it is metaphysically impossible that Christopher does not choose the first options if he is rational and has normal human preferences. It is incompatible with what it is to be rational and with what it is to have the relevant preferences to choose the second option. A person who chooses the second option in the very same

[13] I am not using "psychological states" or "psychological properties" in the way proposed by Chalmers (1997) where the psychological—as opposed to the phenomenal—is defined by causal role. Experiential properties are, as I use the term here, a subclass of psychological properties.

objective situation thereby shows that she has different preferences or lacks rationality. So: relative to an appropriately chosen moment before the decision, there is no metaphysically possible situation fulfilling the three conditions in the definition of psychological determination in which Christopher acts otherwise.

In Daniela's case too we can easily introduce assumptions about her motivation, her beliefs, preferences and value attitudes and about her capacity to judge and to act rationally which make it metaphysically impossible that a person who is psychologically like her votes in the very same objective situation for the Nazi-candidate instead of voting for candidate A. If in a considered counterfactual situation which is like the real situation in all respects mentioned in (a) and (c) in definition 2, Daniela does not votes for candidate B then the considered situation must be one in which Daniela is psychologically different from how she is in the real world (perhaps she has different political convictions, perhaps she has different information, or she has been threatened etc).

I conclude that it is easy to elaborate the above four examples in such a way that they are all examples of psychological determination in the sense defined above. In each of these cases there are moments long before the person acts such that it is psychologically determined relative to those moments that the person acts in the way she does. But aren't those examples all typical cases of free action? Let us go through each of these examples once more with the question about freedom in mind.

Anton decides in the light of the reasons available to him. He does not deliberate because the situation is clear and the act required must be executed immediately. He acts in accordance with his values and convictions; he does exactly what he would have decided to do had he taken time for deliberation. He is praiseworthy for his quick reaction. Anton's act is a typical case of free action in the sense of 'freedom' required for moral responsibility.

I see two potential objections against the claim that Anton's action is to be qualified as free in the relevant sense. Someone may urge that Anton's quick response is triggered by emotion. As a normal and caring person, so someone may argue, Anton is shocked by his perception of danger for the boy and he reacts out of that sudden emotion. Acting immediately out of an uncontrolled sudden emotion, so the opponent may insist, is not a case of free action. If Anton is praiseworthy, so the opponent may continue, this is not for his action or for his decision to act in that way but rather for his character and emotional dispositions.

It may be true that most people who perceive a child in serious danger will immediately react with a strong emotion and will then perhaps act, in some sense to be clarified, out of that emotion. Should we conclude that the act is not free? One should not allow oneself to be led to this conclusion by an inadequate picture of what it is to react out of a sudden emotion. The inadequate picture I have in mind may be described like this: sudden emotions are something that happens to the person. Acts out of a sudden emotion are automatic reactions; they are not under the person's control. I agree that sudden emotions happen to us and are not actively brought about by the person concerned. It does not follow, however, that acts out of emotion are automatic responses. If one fully appreciates what it is to be active one will insist that acts out of sudden emotions are active behavior. The horrified person who immediately acts in order to save the child moves actively and does so with a

specific intention. Furthermore, what she does is not out of control. She has a clear intention and she realizes that intention as quickly and effectively as she can. She does exactly what she considers to be the right thing to do for her in the situation at hand and we may assume her to be fully aware of that fact while she is acting. Strong emotion does not turn the person into a microphysically determined system where no genuine activity occurs and they do not cause the person to execute movements quite independently of her preferences and convictions. To talk of 'automatic responses' in that context is quite beside the point.

There is a tendency to think of strong emotion as 'taking over control' thereby depriving the person of her control over her own behavior. This tendency is surely based on an oversimplified view, however, about the relation between rationality and emotion. I do not wish to deny that emotions sometimes keep us from hesitating when we should hesitate and from deliberating when we should deliberate. And perhaps in extreme cases of strong emotion we cannot be considered responsible for not having hesitated and for not having deliberated and the resulting act might rightly be considered unfree in such cases. But Anton's case, if we assume him to act out of emotion, is not of that kind. He should not hesitate since the situation was clear enough and we may well assume that he is clearly aware of that fact when he immediately jumps up to save the child. Despite the role of emotion in Anton's action we can plausibly attribute the immediacy of his reaction to his rational insight about the objective situation: he just sees without any further reflection that no time must be lost and he acts accordingly. It would have been a bad mistake had Anton started to deliberate about what he should do. No deliberation was necessary to understand what must be done and this fact too can be assumed to be immediately evident to Anton on the basis of his perception of the situation. Anton's perception of the situation rationally justifies his immediate act. If there is a strong emotion of fear for the child involved, then this does not prevent his reaction from being rational and free. If there is such an emotion involved it plausibly helps Anton to correctly perceive the situation in the way he does.

For those who disagree despite these considerations with the claim that acts out of strong emotion can be rational and free in the relevant sense Anton's example can be modified. It is plausible that most people in Anton's situation are terrified by the thought that the child might die; but obviously this is not necessarily so. We can therefore simply stipulate that Anton is not influenced by strong emotion and that he acts on the basis of rational insight. To make this a plausible situation we can add that Anton immediately recognizes that there is no serious risk for the child if he quickly intervenes.

A second objection might be that Anton's quick reaction cannot count as a free action precisely because he reacts so rapidly. He does what he sees must be done but he does it without deliberating about different options. He does not form any preceding intention; he chooses by acting or so one may argue. But should we admit that such rapid reactions cannot be free actions? Does free action require a preceding process of conscious pondering of various alternatives? The answer must be negative. Two intuitions, both misguided, may motivate a positive reply. First, one might think of such quick reactions as 'automatic'. But they are not: they are *actively* brought about by the subject concerned. Second, one might think that free

action requires rationality and that such quick responses cannot count as rational since they do not involve any deliberation. But why should we accept that rational action requires deliberation? Anton does not deliberate since he immediately sees that there is no deliberation required, he immediately sees that to start deliberating would be a mistake and he immediately sees what the result of such deliberation would have to be. In such a situation where it is obvious that immediate action is rationally and morally required, to hesitate in order to think would rather show a lack of rationality. The claim that the absence of deliberation renders the action arational in a way which precludes freedom has, furthermore, a counterintuitive consequence: a less experienced person may not immediately understand the situation and therefore hesitate about what to do. He or she may perhaps call for the parents and wait a moment for their reply thereby losing precious seconds. Anton's quick reaction manifests his superior practical competence and his superior capacity to quickly recognize the relevant features of a given situation. So the claim at issue implies that the action of a less competent person who, contrary to Anton, is unable to quickly recognize the relevant features of the situation would be acting rationally while this cannot be said of Anton in virtue of his superior capacities. Such a result should make one pause. It casts serious doubt on the theoretical assumptions which imply it.

I conclude that Anton's action—when supplemented with appropriate assumptions about its motivation—is free in the relevant sense of freedom. However, as explained earlier, Anton's action is psychologically determined: there are no metaphysically possible counterfactual courses of events in which—in the relevantly same situation concerning external conditions and concerning Anton's psychological properties—Anton does not save the child. In short: it is metaphysically impossible that Anton acts otherwise than he does in the so specified situation. Anton's case is a counterexample to the following simple version of the so-called PAP-principle: for an act to be free it must be possible that the person acts otherwise in the very same situation.

At this point, the following libertarian intuition may immediately come to mind and raise doubts about the result obtained: isn't it obvious that it was up to Anton to remain seated? Isn't it clear that both options where open and up to him? And isn't it obvious that Anton is under the impression that it is within his causal powers to act otherwise? And if this impression is an illusion—as the result obtained seems to suggest—don't we have to say then that Anton did not act freely after all? I share the intuition behind those questions but the conflict between that intuition and the result obtained is only apparent. This is what I would like to explain next.

According to the GAA view, there is a clear sense in which—despite the fact that Anton's act is psychologically determined—all the following claims about Anton are nonetheless correct: the option not to save the child is within Anton's causal powers. He can *actively* move in a way which would be an action of that kind, he knows that he can and he is in a phenomenally manifest way aware of this capacity. But he does not exercise this capacity since the option has no attraction for him at all. There is no reason for him to let the child die. He has no rational, moral and no other motivation for *not* doing what he takes to be the right thing to do in the

situation at hand. Despite the fact that it is within his causal powers not to save the child he does not use those causal powers simply because he has no reason to do so.

Here is a way in which one might interpret what has just been said which is *not* intended: Anton has the bodily capacity to remain seated. However, given his psychological condition it is impossible for him to remain seated in the situation described. It is true that Anton *can* act otherwise only in the following sense: there is a movement which Anton has the general capacity to execute which is, in his particular situation, an action of letting the child die. Given his knowledge about the situation and his other attitudes and values it is, however, impossible for Anton to execute that movement in that concrete situation.[14]

This way to describe Anton's situation is however utterly inadequate according to the GAA view here proposed. Anton not only has the general capacity to move in that particular way, rather he has the causal power to actively bring that movement about in the particular situation in which he is. He is not deprived of that causal power by his awareness of the reasons speaking against that option. He still has the relevant causal power and he is aware of having that causal power. He may at the same time be aware of the fact that he is psychologically determined to save the child in the sense explicated: he may be aware of the fact that—given his convictions and his attitudes towards human life—it is simply impossible *that* he will act otherwise. But the content of this awareness is in no conflict with the plausible simultaneous awareness of Anton that it is possible for him—in the sense of being within his causal powers—not to save the child. A natural way to express this difference in ordinary language is perhaps the formulation just chosen in the previous sentences: it is possible *for* Anton to act otherwise but it is not possible *that* Anton acts otherwise. The libertarian mistake and the truth in the libertarian view can then be stated, briefly, as follows: the libertarian rightly insists that for an agent's act to be free it must be possible *for* the agent to act otherwise. But he mistakenly concludes that, therefore, for an agent's act to be free it must be possible *that* the agent acts otherwise.

But how can there be a difference, as just suggested, between the following claims C1 and C2?

(C1) It is possible that agent A does X in situation S.
and
(C2) It is possible for the agent A to do X in situation S.

According to the GAA view this question has a quite simple answer and the difference is not mysterious as it might seem at first sight. (C1) is a claim about metaphysical possibilities and thereby a claim about which actions are and which actions are not compatible with the nature of the psychological states of person A in situation S. (C2) is a claim about the causal powers of the person concerned in the situation S to bring about movements which constitute various actions in the situation at hand. The difference between (C1) and (C2) is due to the fact that psychological properties, in general, exclude courses of action as impossible in a way which has nothing to do with the restriction of causal powers. It is impossible

[14] An analysis along these lines is proposed in Baierlé (2016).

that Anton does not save the child because Anton happens to have a perception of the situation, attitudes towards the value of life and beliefs about what he can do in the situation at hand which make it obvious for him that jumping up in order to save the child is the thing to do. He is therefore motivated to do so and he has no motivation not to do so; being a rational agent he uses his causal power to bring about a series of movements which is—in those circumstances—a saving of the child. But the reasons available to Anton influence his behavior only *in virtue* of his capacity to actively bring about those movements (with the relevant intention in the light of those reasons). Anton's reasons do *not* by themselves cause the movement of Anton's limbs, nor do any mental events (supposedly the recognition of those reasons) cause those movements. It is Anton who makes happen (actively brings about) what happens, not any events happening inside his body or brain. Despite all those things happening inside his body or brain and despite his awareness of the reasons speaking in favor of jumping up to save the child, Anton is not deprived of his causal power to behave and thereby act otherwise; he still has that power, he simply doesn't act otherwise because he is rational and has no reason to act otherwise; therefore, he does not act otherwise in any metaphysically possible situation in which he has those psychological properties since those properties are incompatible by their nature with not trying to save the child.

Given the above discussion of Anton's case it will be quite obvious how the other examples can be used within the GAA view to argue against the libertarian thesis that free action requires the possibility that the person could have acted otherwise. Beatrice acts in accordance with her long term decisions about the kind of life she wishes to lead or so we may assume. She is in no way forced to go to school on every Monday in order to teach mathematics. To do so is what she wants to do. Intuitively it would be odd to say that she does not act freely. Someone whose theoretical account of freedom implies that Beatrice is not free in those actions might try to save his or her theory arguing that Beatrice is just acting out of habit and actions done out of habit are not free. Again (like in the case of emotion), the idea that actions done by habit are not free may have to do with the thought that they are automatic. But here again the GAA theorist will insist that to think so is a mistake. Actions done by habit are still active behavior and so they should not be qualified as automatic. Another way to argue that Beatrice does not act freely would be to say that actions done out of habit cannot be free because they are arational since done without previous deliberation.

But is there any good reason to suppose that previous deliberation is required for an act to be free? There does not seem to be any such reason. To see this let us suppose that Beatrice is aware of the reasons speaking strongly in favor of going to school in order to teach. Even if she acts out of habit those reasons may well be present to her each Monday morning when she prepares herself to leave the house. She may well be aware of having accepted that duty and she may be looking forward to teaching. She executes the complex behavior of preparing herself for the lesson and of leaving in order to go to school in the light of those reasons which, as we may assume, are present to her while she acts in a phenomenally manifest manner. Since those reasons are so compelling, she does not even start to consider other options such as staying at home on her sunny balcony in order to read an

interesting book. It would be odd to assume that in order for her act to be free she must have doubts about what to do. Compare Beatrice with her colleague Beat who is much less committed to his duties, who dislikes teaching and who is otherwise in a comparable situation. Each Monday morning he hesitates a moment while still lying in bed about whether he should go to school or rather take a day off. Only after a few minutes of pondering the various options open to him he finally decides to go teaching. I cannot see any plausibility in the assumption that Beat is free in his action while Beatrice is not free in hers in virtue of the difference between them which is that Beat has to think about what to do while Beatrice knows it immediately and therefore can skip the deliberation process.

All these considerations speak in favor of the claim that Beatrice is acting freely in the relevant sense when she goes to school every Monday morning in order to teach mathematics. But every friend who knows Beatrice well enough can predict with certainty that Beatrice will act in that manner. And again it seems clear that the very nature of Beatrice's psychological states is incompatible with a different behavior under the circumstances described. So Beatrice's action is psychologically determined and we must conclude—contrary to the libertarian view—that psychological determination does not exclude freedom.

It has been argued against the libertarian view that the freedom libertarians claim people to enjoy is nothing but the capacity to act against reason. The argument may be put as follows: the libertarian insists that, for a rational choice to be free, it must be possible for the agent to act otherwise and so, in that case, against reason. But if this is what distinguishes (among rational actions) those that are free from those that are not free then it is hard to see why freedom should be valuable.[15] In order to respond to this objection the libertarian must insist that there is more to libertarian freedom than just the capacity to act otherwise. But if he wants to insist that a person who acts rationally could have acted otherwise *under the very same circumstances*, then it is difficult to see how he can answer the challenge.

The GAA view can avoid this problem in a way which does duty to the libertarian basic intuition. The GAA view rejects the libertarian claim that a choice is free only if there was an *open possibility* that the person chooses otherwise and yet preserves the intuitive core of the libertarian intuition that a rational act is free *only if* the person has the *capacity* to act otherwise. That capacity must be interpreted, according to the GAA view, in the following way: it is *within the person's causal powers* to act otherwise. According to the GAA view there is no contradiction between (a) it being within the causal powers of a person to act otherwise in given circumstances and (b) it being psychologically and metaphysically excluded that the person uses that capacity in those circumstances.[16]

[15] Such an argument may be found in Wolf (1990, p. 57). On the basis of the reasoning just mentioned she asks with respect to libertarian freedom: "Why should one want an ability one never wants to exercise? Why should one mind if (…) one is *inescapably* sane?" Under the title "the rational cost problem" this difficulty for the libertarian is discussed in detail in Michael Garrett (2013).

[16] One might think that the objection here at issue can be reformulated against the view just sketched saying that there is no value in having the causal powers to act against reason since it is a bad thing to make use of that power. But this new version of the argument has no force against the GAA view for the following reasons: First, according to that view, having the causal powers to act in various ways in a

Let us finally turn to the examples involving Christopher and Daniela. A theorist who wishes to defend the view that freedom excludes psychological determination might try once again to undermine the thesis that Christopher and Daniela act freely by suggesting that they are in some way compelled to act as they do. Without going into any details about such a further description of the two examples it is clear, however, that both examples can be supplemented in a way such that no potentially freedom undermining factors play any role. Both, Christopher and Daniela decide, or so we can stipulate, in a non-emotional and well-reflected manner on the basis of well-founded information. Given their respective preferences and beliefs and given their respective rationality it is once again predictable for a well-informed observer how they will decide and it is incompatible with the very nature of the psychological properties they instantiate that they act differently under the very same conditions. Both examples are well suited to illustrate the general point: in any situation in which one of the options available to a given rational agent is clearly preferable (given the preferences and beliefs and all other relevant psychological properties of the person concerned) to all other available options, it is metaphysically excluded that the person will act otherwise than she does under exactly those circumstances. The thesis that freedom excludes psychological determination therefore implies that the choice of the rationally best alternative is never free if the chosen alternative is *clearly* the best one in the light of the reasons available to the person. This consequence strikes me as unacceptable. Since there does not seem to be any way to deny that such choices are psychologically determined and yet typical cases of free choice the libertarian idea that freedom implies the possibility *that* the person acts otherwise under the very same circumstances must be abandoned. According to the GAA view the intuitive tendency to nonetheless insist on the contrary thesis is due to a confusion between what is possible *for* an agent to do (what is within the agent's causal powers in the concrete circumstances at issue) and what is possible *that* the person does in those circumstances (what is compatible with the nature of her psychological properties).

So far it may seem as if the GAA view is friendlier to the compatibilist than to the libertarian. But this impression will vanish once one considers what the proponent of the GAA view has to say about easy choices and about tough choices. Let us start with the latter.

9 Tough Choices and Easy Choices

Sophie's choice in the film and in the novel "Sophie's choice" is a particularly telling and horrible example of a *tough* choice.[17] Sophie is forced by members of the Gestapo to decide who of her two children, an older daughter and a younger son,

Footnote 16 continued
given situation is a general feature of action and does not distinguish free action from unfree action. Second, the proposal makes intelligible the sense in which an act is up to the agent (in a sense necessary for acts in general and free acts in particular) and 'being up to the agent' in *that* sense is arguably a necessary ingredient of valuable freedom.

[17] The film *Sophie's Choice* with Meryl Streep in the role of Sophie came out in 1982 and was directed by Alan J. Pakula, who adapted William Styron's homonymous novel.

will be murdered and who of them will survive. She saves the boy. For the purposes of the discussion which follows let us consider the example as if it was actual and not just fictional.

According to the compatibilist standard view, if Sophie decides to save the boy, then it must be possible, in principle, to explain why she does so. Perhaps her love for the boy is deeper or his younger age motivates her choice. According to the standard compatibilist view a contrasting explanation must in principle be available, in other words: if, in a considered counterfactual situation, Sophie decides to save the daughter rather than the boy, then—if all external conditions are alike—there must be a difference concerning Sophie, presumably concerning her psychological properties, that explains why she acts differently in that counterfactual situation. The libertarian disagrees. According to the libertarian, Sophie's choice need not be psychologically or otherwise determined.[18] The libertarian will typically say that the case might well be such that it is possible (in all relevant senses: nomologically, psychologically and metaphysically possible) *that* Sophie saves the girl in exactly the same situation.

With respect to Sophie's case and other tough choices, the GAA view sides with the libertarian. Using the above distinction between 'possible that' and 'possible for' we may say this: within the GAA view the claim is well motivated that Sophie's situation *can* be such that it is possible *for her* to save the girl (it is within her causal powers) *and* it is possible *that* she saves the girl in exactly the same situation with respect to all relevant factors (it is metaphysically possible, compatible in particular with the nature of her psychological states) that she chooses otherwise in the very same situation.

Within the GAA view it is not surprising and not mysterious if a person who does not see any reason to prefer one of two available options takes an *arbitrary* choice just by *actively* taking a decision or by actively realizing the behavior which is the realization of one possible act rather than of the other. Sophie is aware that she must take a decision. She has, let us suppose, no preference with respect to these two horrible options. Therefore, at some point, she just reaches out for the boy and thereby communicates how she decided. It was possible *that* she reaches out for the girl and it was possible *for her* to reach out for the girl. In metaphysically and nomologically possible circumstances she arbitrarily reaches out for the girl. The objection that there *must* be a relevant psychological difference between these two cases which explains why she saves the daughter in the counterfactual situation and the son in the real world has no force against the GAA theorist. The understanding proposed by the GAA theorist of what it is to be active in one's behavior precisely includes the subject's capacity to *arbitrarily* realize one of two open options in a case where reason in a broad sense which includes emotional experience, attitude and perception delivers no verdict about what to do. To put it in more abstract terms: the GAA view proposes an account of what it is to be active in one's behavior according to which genuinely arbitrary choices are possible; and genuinely

[18] The libertarian will deny that Sophie's choice is nomologically determined in the usual sense: it cannot be deduced from adequately chosen previous conditions plus the laws of nature that the choses (or acts) as she does.

arbitrary choices do *not* require contrasting explanations (this is in fact what it is for a choice to be genuinely arbitrary).[19] Therefore, to argue against the GAA view using the assumption that contrasting explanations of different courses of action must always be available is to beg the question against the GAA view.

One might be tempted to think that the compatibilist is right and the libertarian wrong about Sophie's case having in mind the following reason: it is hard to believe that a person in such a serious and deeply disturbing situation would take a genuinely arbitrary choice. Humans in such serious situations will rather bring themselves into a state of mind where one of the possible choices appears in a better light than the other. But to reject the libertarian thesis about the case for that reason would be to interpret the libertarian in an unnecessarily strong and uncharitable way. The libertarian need not and should not insist on the possibility that the choice is psychologically underdetermined until the very last moment. He or she should rather defend a weaker thesis: there are moments m before the choice such that with respect to m the act is not psychologically determined in the sense of the above proposed definition. To say this of Sophie, for instance, about a moment a few seconds before she reaches out for the boy does not exclude that Sophie actively sets up her mind between that moment and her later choice thereby arbitrarily giving, for instance, more weight to one kind of consideration than to another. In such a case it may be that Sophie's choice is not psychologically determined at a certain moment m a few seconds before she acts and yet *is* psychologically determined a little later as a result of Sophie's own active and not psychologically or otherwise determined intervention.

Perhaps *tough* choices are seldom or never taken by just arbitrarily choosing 'in the last moment' by actively realizing one of the options available. But there are other choices, *easy choices*, where it is quite plausible to assume that this is often precisely what happens. Easy choices concern different options in cases where nothing of any importance depends on which option is realized. Suppose you like ginger ice cream and you also like lemon ice cream and you have no preference. At some point when ordering ice you must decide and you choose in a genuinely arbitrary manner. You utter "ginger" instead of "lemon" at the right moment.

There *can* be an explanation of your choice. Perhaps you decide as you do due to some psychological factor you are unaware of. The GAA view only insists that there *need* not be such an explanation and, as noted earlier, the view that there need not be such an explanation is well-motivated within the GAA view. I take this element of the GAA view to speak in its favor. We often experience our own choices as genuinely arbitrary. Perhaps this is, in many cases, an illusion. However, if a theory implies that the impression of arbitrary choices is illusionary without any exception and even necessarily so then this provides reason to reject that theory.

These considerations about tough and easy choices complete my argument for the following claim: the GAA view integrates the true core of the compatibilist theory and the true core of the libertarian theory while avoiding the weaknesses of both. In this sense the GAA view provides a 'synthesis' of the two competing theses. If the synthesis here proposed can be successfully defended within the GAA view and if

[19] For further elaboration of this point compare Nida-Rümelin (in preparation) chapter 8.

no alternative to the GAA view can provide a similar synthesis, then this provides a strong argument for the claim that the GAA view is correct: conscious animals are genuinely active; they are not microphysically determined in their behavior.

10 Concluding Remark

In the present paper I briefly sketched a phenomenological argument and I elaborated an indirect theoretical argument in favor of the GAA view. But even the best development of these arguments are unlikely to convince those who reject the GAA view for reasons based on what they take to be established by empirical science. Further defending the GAA view will have to involve answering a number of scientifically motivated doubts and objections. Such an argument will have to show:

1. There is at present no well-established scientific knowledge which excludes the truth of the GAA view.
2. Possible future scientific development may lead to empirical results speaking in favor of the GAA view.

The philosophical opponent is right in claiming that any argument for the GAA view remains unsatisfying and incomplete if not supplemented by convincing reasons in favor of (1) and, even better, in favor of (2) as well.

But there is a similar scientifically motivated objection that should be seen to be based on unjustified expectations. The objection I have in mind might be put like this: we have no idea of *how* and *where* the conscious subject might actively intervene on the level of physical processes happening in the brain. As long as no plausible model of how that might work is provided the GAA view (just like the ACV) need not be taken seriously. This kind of objection is, however, ill-founded. It should be obvious that such a model cannot be developed without detailed knowledge about the brain. A philosopher should remain silent about how that might work and must leave the development of such a theory to empirical science.[20]

References

Baierlé, E. (2016). *The phenomenology of choice*. Ph.D. Thesis. Fribourg: University of Fribourg.

Beckermann, A. (2006). Neuronale Determiniertheit und Freiheit. In K. Köchy & D. Stederoth (Eds.), *Willensfreiheit als interdisziplinäres Problem* (pp. 289–304). Freiburg: Alber.

Broad, C. D. (1953). Determinism, indeterminism and libertarianism. In C. D. Broad (Ed.), *Ethics and the history of philosophy* (pp. 195–217). London: Routledge (Reprinted by Routledge in 2014).

[20] The view here presented has in great part developed in discussion with my collaborators Emmanuel Baierlé, Franziska Müller and Jacob Naito in the context of our research project "Philosophy and Phenomenology of Agency" (1.10.2010–31.1.2014, project reference: PDFMP1-132455). I would like to thank the Swiss National Science Foundation for its support which gave us the very gratifying opportunity to pursue these issues within philosophy of agency over several years in personal research, in the context of regular group meetings and in international conferences.

Chalmers, D. (1997). *The conscious mind*. Oxford: Oxford University Press.
Chisholm, R. M. (1964). Human freedom and the self. In R. Kane (Ed.), *Free will*. Oxford: Blackwell.
Fischer, J. M., Kane, R., Pereboom, D., & Vargas, M. (2007), *Four views on free will*, *great debates in philosophy*, Series editor: Ernest Sosa. Blackwell.
Garrett, M. (2013). Freedom and unpredictability. *Inquiry, 56*(6), 666–680.
Horgan, T. (2007). Agentive phenomenal intentionality and the limits of introspection. *Psyche, 31*(1), 1–29.
Horgan, T., Tienson, J., & Graham, G. (2003). The phenomenology of first person agency. In S. Walter & H.-D. Heckmann (Eds.), *Physicalism and mental causation. The metaphysics of mind and action*. Exeter: Imprint Academics.
Keil, G. (2014). *Willensfreiheit und Determinismus*. Stuttgart: Reclam Taschenbuch.
Nida-Rümelin, J. (2005). *Über menschliche Freiheit*. Stuttgart: Reclam.
Nida-Rümelin, M. (2007). Doings and subject causation. *Erkenntnis, 67*(2), 255–272. (Special issue edited by A. Newen, V. Hoffmann, M. Esfeld, Mental Causation, Externalism, and Self-Knowledge).
Nida-Rümelin, M. (2016a). The experience property framework—A misleading paradigm. *Synthese* (special issue: *Trends in Philosophy of Language and Mind*, edited by Hanoch Ben-Yami, pp. 1–27).
Nida-Rümelin, M. (2016b). Active animals and human freedom. In H. Rott et al. (Ed.), *Festschrift for Wolfgang Spohn* (pp. 339–378). Münster: Mentis.
Nida-Rümelin, M. (in preparation). *Conscious individuals. Sketch of a theory*. Oxford University Press (under contract with).
O'Connor, T. (1996). Why agent causation? *Philosophical Topics, 24,* 143–158.
O'Connor, T. (2002/2010). Free will. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Fall 2014 Edn) <http://plato.stanford.edu/archives/fall2014/entries/freewill/>.
Pauen, M. (2004). *Illusion Freiheit?* Frankfurt am Main: S. Fischer Verlage.
Siegel, S. (2012). *The contents of visual experience*. Oxford: Oxford University Press.
Spohn, W. (2012). Der Kern der Willensfreiheit. In D. Sturma (Ed.), *Vernunft und Freiheit*: *Zur praktischen Philosophie von Julian Nida-Rümelin* (pp. 71–90). Berlin: De Gruyter.
Steward, H. (2012a). The metaphysical presuppositions of moral responsibility. *The Journal of Ethics, 16*(2), 241–271.
Steward, H. (2012b). *A metaphysics for freedom*. Oxford: Oxford University Press.
Wolf, S. (1990). *Freedom within reason*. Oxford: Oxford University Press.