

**ENACTING THE MORAL SELF  
– COMBINING ENACTIVIST COGNITIVE SCIENCE  
WITH MEAD’S PRAGMATISM**

**Martin Weichold**

*University of Regensburg, Germany*

[Martin.Weichold@ur.de](mailto:Martin.Weichold@ur.de)

**ABSTRACT:** This paper suggests teaming up enactivist cognitive science with George Herbert Mead’s pragmatist conception of social self-constitution, in order to develop a new account of moral agency. Such a new account is necessary because our everyday conception of being autonomous individuals, or moral selves, is seriously challenged by a wealth of empirical findings from the cognitive and social sciences. But while those findings challenge the idea that there is a real, pre-social autonomous self who is the subject of human actions, the findings leave conceptual space for a Meadian social enactivist account of the self as an object, that is, as an objectified *self-conception*, developed by human animals in the course of embodied social interactions. This objectified self-conception can then, it is argued, feed back into our embodied social interactions, so that we humans actually play our selves – play our roles as pre-social autonomous individuals.

## 1. Introduction

I am an autonomous individual: I can think for myself, can see that some of the social norms we currently have are morally wrong, can make free and independent decisions which authentically reflect my true inner values, and am responsible for the actions of my earlier and future selves. Many people, philosophers or otherwise, would assuredly claim that the same statement holds true of them as well. Some might even regard the statement as a simple truism. However, this apparent truism seems to fly in the face of many views in recent cognitive and social sciences – and in particular, it seems to fly in the face of George Herbert Mead’s pragmatist account of the social constitution of the self, and it seems to fly in the face of the recently very popular enactivist cognitive science. For, the apparent truism presupposes that we humans are autonomous moral agents who can criticize the social order from outside, in virtue of detached moral reasoning. However, according to Mead, we humans only become persons with selves by being constantly interwoven into social

interactions – then, humans are not external judges of social norms, but rather products of the social order. In Mead’s words, we should “assume a social process or social order as the logical and biological precondition of the appearance of the selves of the individual organisms involved in that process or belonging to that order” (Mead 1934, 222). Enactivism suggests that at least “the vast sea of what humans do and experience is best understood by appealing to dynamically unfolding, situated embodied interactions and engagements with worldly offerings” (Hutto & Myin 2013, ix) – thereby, it seems, denying the existence of disembodied autonomous agents who act in virtue of considering thoughts with moral content. Thus, at least at first glance it seems that a simple truism about moral selfhood gives rise to a complicated problem for Meadian pragmatism and enactivist cognitive science.

In this paper, however, I will argue that it is actually the other way around: The apparent truism that we human beings are pre-social autonomous individuals, or moral selves, is actually brought into severe danger by a wealth of recent empirical findings from both the cognitive and the social sciences. But if the idea that we are autonomous individuals were just an illusion, the consequences would be disastrous: For what would happen to the criticism of social norms, authentic actions, rights and duties, and responsibility for past and future actions? I will suggest that there is a way to avoid this disastrous situation: enactivist cognitive science and Mead’s pragmatist account of social self-constitution can be teamed up in order to construe a new account of moral selfhood.

According to this suggestion, human agency, as a biological and psychological phenomenon, really consists in “dynamically unfolding, situated embodied interactions and engagements with worldly offerings” (Hutto & Myin 2013, ix), without there being an autonomous self who is the subject of the action. But in virtue of being constantly engaged in social interactions, human animals develop a *self-conception*, a conception of their selves as objects – as things which can weight and measured, looked at and thought about. And part of

this objectified self-conception is that humans are autonomous individuals. This detached self-conception then – it is suggested – feeds back into our embodied social interactions with others. And in such a way, we humans can, so to speak, play the role of an autonomous individual, even if we are not autonomous individuals in reality. If this proposal is on the right track, it would not only show that the idea that we are autonomous individuals is not in conflict with enactivism and Mead’s pragmatism – enactivism and Mead would show how moral selfhood becomes possible in the first place.

By making this proposal, I am of course also making a suggestion for how “4e cognitive science” and classical pragmatism might be fruitfully combined. Combining enactivist cognitive science and Mead’s pragmatist account of social self-constitution is particularly fruitful, I suggest, not only because enactivism and Mead share so much, but also because it promises to show how the reach of enactivism can be extended to the social domain, and this in a novel way (different, for instance, from the so-called participatory sense-making approach, see e.g. de Jaegher & di Paolo 2007). Indeed, it would be appropriate to call Mead a “social enactivist”. Yet the philosophical depth, innovativeness, and plausibility of enactivism and Mead’s pragmatism can be seen best once their accounts are brought into conversation with mainstream views – and this is why this paper aims to set up a dialogue between enactivism, Mead, and classical moral philosophy.

The paper unfolds as follows. In the next section (section 2), I will explain the central problem in more detail, contrasting our everyday self-conception as autonomous individuals with recent findings from the cognitive and social sciences. Having set up the central problem, section 3 departs from the main line of argumentation for a moment in order to refute a view I call “harmonism” – that is, the idea that there actually is no tension in our human self-understanding, because things like autonomy and free will can be unproblematically conceived of as part of the natural world. It is worth showing the implausibility of

harmonism since it can be arrived at through a misunderstanding of anti-dualism, which is very prominent amongst pragmatist thinkers. Section 4 returns to the main line of argumentation and introduces enactivism as one crucial element for a new explanation for the emergence of moral selfhood. But since enactivism cannot explain moral selfhood on its own, section 5 introduces Mead’s account of social self-constitution as a further building block for approaching moral selfhood in a new way. Even though this can already explain a great deal, the enactivist, constructivist, constitutionist spirit of both enactivism and Mead’s account of selfhood is in tension with the realism about autonomous individuals which is part of our everyday moral practices. This is why section 6 introduces a further and final building block, combining enactivism and Mead in a more sophisticated way. Finally, section 7 shows how this new account of moral selfhood allows us to conceive of human nature as unitary, but also to appreciate the tensions and disharmonies between our biological nature and our social self-conceptions – disharmonies which make us the beings we are.

## **2. Human Nature in Inner Conflict**

There are many ways of thinking of human nature, ranging from the conceptions of world religions to the very diverse views in antiquity, enlightenment, and evolutionary theory. But I want to suggest that today two views of ourselves are of particular importance to us, views which are both indispensable, but which are in conflict with each other. One view is the conception of ourselves as moral agents, a conception which we presuppose in our everyday moral practices. The other view is the idea that we human beings are essentially animals, habitually interacting with our material and social environments in a bodily way, not too dissimilar to the way other animals behave – a conception suggested both by recent cognitive and social science. – Let me first spell out our everyday self-understanding in more detail,

and then show how it can be seen as challenged by recent empirical findings. These expositions will be more abstract and general, but they are necessary for setting the stage for the new combination of enactivist cognitive science and Mead's pragmatism which is to follow.

As I have already alluded to in the beginning of the paper, in our everyday moral practices we think of ourselves as active beings who can freely act in the light of reasons, can criticize social norms, and are responsible for their deeds. Of course, this is often a rather unarticulated and implicit conception, but it clearly guides our everyday behavior, and so it is worth highlighting some aspects of it.

One aspect of our everyday self-conception is that we think that we are agents, that we can do things such as buying biscuits or blaming others. In our everyday practice, we clearly ascribe some events to ourselves as their authors, saying "that was me", or "it was not me". This conception is of high practical importance because it allows us to take responsibility for certain events. In philosophy, this conception is represented in many action theories (e.g. Anscombe 1963, Davidson 1963).

Another related set of features of our everyday self-conception is that we have full control over our behavior, and act with free will and autonomy. Clearly we say things such as "Sorry, but I couldn't control myself", or "He is not a very autonomous and independent person but just does what his parents expect him to do". This conception is of practical importance for taking responsibility and for excluding someone from responsibility. In philosophy, it can be found, amongst other places, in discussions about free will (e.g. Kane 2007, Fischer & Ravizza 1998).

We also think that we are able to see reasons and act on them, in particular moral reasons. For instance, we say things such as "What were your reasons?", or "It was not rational, I just reacted emotionally". This idea is practically important because it allows us to judge actions, and to sometimes guide our behavior. This part of our self-conception is represented in philosophical

accounts of rationality and reasons (e.g. Parfit 2011, Korsgaard 2009).

Finally, another important set of ideas entailed in our everyday self-conception is that we are particular persons with a distinctive personality, identity, character, and true self. For example, we say things like "She is not the same person anymore", "he has a good character", or "I was not really myself". This set of ideas is of high practical importance because it shapes how we act in the present, opens up the possibility to relate to our past and future selves, and influences how we make sense of others and interact with them. In philosophy, these ideas are represented in conceptions of the self (e.g. Gallagher 2011), in ancient and modern conceptions of character (e.g. Aristotle, NE), and in discussion about personal identity (e.g. Parfit 1984).

Much more can be said about the particular aspects of our everyday self-conception – but for the moment, it is sufficient to notice how rich and multi-faceted our everyday self-conception is, and how important it is for our everyday practices. Here is the reason why it is crucial to be aware of the multifaceted nature and the practical importance of our self-conception: Some philosophers assume that all there is to empirical challenges of our everyday self-understanding are neuroscientifically informed arguments against incompatibilist theories of free will. But this would be wrong in at least three ways: First, it is not only abstract theories, but our very practical everyday self-conception which is in danger. Second, the idea that we have free will is only one of many different aspects of our everyday self-conception. And third, as it will be made clear shortly, it is not only neuroscience which challenges our everyday self-conception, but also findings from other cognitive sciences such as cognitive psychology and social psychology, findings from the social sciences, and even careful philosophical descriptions of phenomena of life.

Indeed, it should not be too surprising if our everyday self-conception would be refuted by cognitive and social science. After all, the core of our everyday

view can already be found in early antiquity, e.g. in the works of Plato and Aristotle. And many other aspects have been developed through Christianity. Yet most of the views dating from that time, on physics or biology for instance, have been replaced today. So one might wish to take serious the possibility that our ancient-old self-conception has to be reconsidered in light of what has been found in controlled, rigorous scientific experiments (cf. Doris 2002).<sup>1</sup>

And there is indeed a wealth of empirical findings which have been used for challenging different aspects of our everyday self-understanding. Given that there are so many different sets of findings from the cognitive and social sciences, I will mainly focus on selected sets of findings from the cognitive sciences during the next paragraphs. The most prominent set of findings is the

---

<sup>1</sup> Of course, this does not mean that, say, Aristotle’s philosophy is as outdated as his physics. Quite to the contrary, Aristotle might have had an outstanding sense for analyzing important phenomena of human life, a sense which many thinkers today are lacking. And it is of course also true that we today might fail to see important phenomena clearly, just because we are viewing them through the dyed lens of modern conceptual frameworks which have become invisible to us, but which shape our outlook on the world in one-sided ways – materialistic, individualistic, or capitalistic ways of thinking might be examples. In these senses, I agree that it can be helpful to return to philosophical views dating back to the time of Plato and Aristotle. Yet I insist that those views have to be checked against findings from recent cognitive and social science. In reluctance to this view, one might argue that so far nothing has really made us change our classical self-conception. So, did the ancient Greeks get it right about human agency after all? But it does not seem very likely that societies which even got the simpler things like biology wrong should be exactly right about the more complicated matters such as cognition in action, which is the topic of sophisticated research e.g. in cognitive psychology and cognitive neuroscience. Rather, I contend, we should distinguish between truth and success conditions. A self-conception can be very successful even if it is wrong. For example, if I believe that I am God’s Chosen One, this might uplift my view on life and make me bear misfortunes more easily – still, this practically successful self-conception would be superstitious. Or, to have a slightly different example: a society which systematically punishes those who deviate from its social norms might be more stable and successful – whether or not things like ultimate desert, true responsibility, or libertarian free will really exist.

influence of unconscious factors on human behavior. Human behavior has been found to be often effectuated by factors humans are not aware of, but which can be successfully used for predicting the behavior (Wegner 2002).<sup>2</sup> Moreover, it has been found that human behavior often proceeds “automatically” and habitually, without being guided by conscious intentions (Bargh & Chartrand 1999). If one is honest to oneself, I submit, one might also see the influence of those unconscious factors in one’s own life, or in the behavior of one’s friends. However, those findings have been used as material for arguments to the effect that we do not have the free will, or the control over our actions, which we take ourselves to possess (Wegner 2002). And similarly, it can be argued that we are not the rational creatures we think we are, insofar as we often do not act in the light of reasons (Sie & Wouters 2010). Moreover, the findings can be utilized for arguing that we are not the self-transparent, self-conscious selves, or agents, which we take ourselves to be (Wilson 2004, Doris 2015).

In our everyday moral practices we think that action is a matter of our conscious control, of our intentions.<sup>3</sup> However, there are many empirical findings suggesting that other factors might have an even more profound effect on the course of human behavior. For instance, it has been found that how well humans can control themselves and withstand temptations is depended upon their current energy – that is, on their glucose level in their blood (Baumeister et al. 1998). And it has been

---

<sup>2</sup> Obviously, the unconscious has already been famously discussed by Freud. However, his theory has been regarded as problematic from a psychological point of view, because it postulates Homunculi (the id, the ego, and the superego), allows for no concrete predications of behavior, and appears to be unfalsifiable (cf. Uleman 2007, 4f.). However, the situation has dramatically changed during the last decades, and now there is a wealth of non-Freudian empirical studies of the unconscious which do meet scientific standards. Expressing this trend, the title of a recent anthology on the topic is “The New Unconscious” (Hassin et al. 2007).

<sup>3</sup> Of course, philosophers in the traditions of Hume and Davidson would emphasize that we have no direct control over our desires, in the service of which we allegedly get to action.

found that human behavior is strongly shaped by moods, habits, and the body (e.g. Gallagher 2005, Pacherie 2011). And I again, I submit, everyone who has more or less carefully observed her own past behavior or the behavior of her friends can know these things from her own experience. However, these findings can be utilized for arguing that our behavior is not controlled by ourselves in the way we think it is.

According to our everyday self-conception, it is primarily “inner” factors which shape our actions, be it our character traits, conscious control, deliberate intentions, or reflective decisions. However, there is a large class of findings, stemming primarily from situationist social psychology, according to which human behavior is essentially shaped by “outer” factors of a given situation (keeping in mind that the dichotomy between “inner” and “outer” is something pragmatists and enactivists might reject). It has been found that situational factors such as finding a dime (Isen & Levin 1972), being in a hurry (Darley & Batson 1973), and playing roles (Zimbardo et al. 1973) tremendously influence behavior, e.g. whether people help others. Famously, Stanley Milgram even found that the situational factor of being in the presence of an authority figure tremendously influences whether normal people are ready to torture an innocent person (Milgram 1974): In what was presented as an experiment on the effect of punishment on learning, normal people gave electric shocks up to 450 Volts to their vehemently protesting fellows, if an authoritative scientist insisted that this has to be done. Admittedly, the amount of influence situational factors have on human behavior by far exceeds what is intuitive – for instance, laymen and professional psychiatrists predicted that between 0.125% and 2% of the persons in the Milgram experiment would give 450 volt electric shocks to a protesting victim, whereas in fact 65% of the persons (in one version) did so. These findings have been used to argue that we humans do not possess classically conceived character traits, since it might, for example, depend more on the situation, and our mood and stress

level whether or not we help another person, rather than on our benevolence or good intentions (Harman 1999, Doris 2002).

Another set of findings concerns the idea that we are creatures who normally act in the light of reasons. It has been found that most of the good reasons we mention in favor of our actions are not statements of our initial motivations, but are rather post-hoc rationalizations uttered for the purpose of justifying ourselves (Nisbett & Wilson 1977, Haidt 2001). Again, any attentive and honest person might be able to detect at least some of those rationalizations in her own behavior and in the behavior of her friends. However, these findings call into doubt our self-conception as agents who normally act in the light of reasons (Sie & Wouters 2010).

A last and heterogeneous set of thoughts calls into question the idea that we are autonomous persons with a unified personality. To begin with, the idea that humans have consistent personality traits has been questioned by findings which suggest that personality traits are at best situation sensitive – the person who is honest with her parents might be dishonest with her friends (Mischel 1968). Moreover, inspired by different findings from social science, many social theories have given up the idea that humans are autonomous selves, focusing instead on the social power mechanisms which shape our constructions of our identities. And even in the recent interdisciplinary debate about the nature of the self, basically no one defends the idea that humans are just by nature persons with agential powers: either the existence of the self is denied (Metzinger 2003, Albahari 2006), or understood as a bodily phenomenon (Zahavi 2005), or as a narrative construction (Dennett 1992, Schechtman 1996, 2014). Those findings and ideas can be used to shed doubt on our everyday self-conceptions as particular persons with distinct personalities and true inner selves.

We must, then, come to the following conclusion: There is an important new problem concerning human nature. Our human self-understanding is terribly conflicted. On the one hand, we conceptualize ourselves

as autonomous individuals. And this conception seems to be the condition of the possibility of the moral practices which enable us to live the lives we live. However, there is an abundance of findings from the cognitive and social sciences which suggest that we are at best human animals with fragmented minds and no inner selves, constantly interacting with a material and social environment against the background of factors such as moods, bodies, habits, energy, and stress-levels. And attentive persons can verify this by observing their own existence. So it seems that we humans are both: autonomous agents and automatic animals. And yet, the conceptions seem to be mutually exclusive. So, what are we?<sup>4</sup>

I would insist that the problem just described is new. However, at first glance one might have a different impression. One might think that the conflict described is actually nothing but a particular version of what Wilfried Sellars (1963) has called the clash between the manifest image of men in the world – our everyday self-understanding –, and the scientific image. According to Sellars, our everyday view on the world might clash with science’s view on the world. However, the problem I described is different from the problem described by Sellars. For, the view in favor of which I argued is opposed to our everyday self-conception is not exactly a scientific image. Indeed, the view that we humans are

---

<sup>4</sup> Obviously, the existence of this problem relies on the credibility of the experiments alluded to, and the plausibility of the arguments based on the findings. But it has become popular amongst some philosophers to “refute” scientific findings and their interpretations, targeting in particular neuroscientifically based arguments against free will (cf. e.g. Mele 2009). However, it has to be pointed to the fact that the mentioned findings stem from a large variety of research programs, and that there are hundreds of studies backing up one result (cf. Doris 2002). Moreover, the results mentioned also resonates with careful observations of existential phenomena, and also with philosophical descriptions of them to be found, e.g., in the works of Dewey (1925), Heidegger (1927), Merleau-Ponty (1942), and Sartre (1943). Consequently, the strategy of “refuting” single findings by re-interpreting single experiments does not seem very promising. But I agree that it would be important to have a closer look at how the mentioned findings conflict with our everyday self-conception.

human animals which habitually interact with their material and social environments is very well supported by recent findings in the cognitive and social sciences. But it is not exactly a scientific view: rather, it is a view which is described on personal (and not subpersonal neuronal) terms, which is related to and mirrored in our everyday existential experience, and which has also been described by philosophers and poets.

But one might think that the problem I have described is not new because it only reflects old dualistic conceptions of human nature. According to the views of Plato and Aristotle, for instance, we humans have different, higher and lower parts of the soul.<sup>5</sup> According to Kant (1787), we are beings which live both in the phenomenal and in the noumenal realm. It is thus tempting to assume that the idea that we are autonomous selves with reason and free will is somehow to be identified with the higher part of the soul, or our existence in the noumenal realm, while our habitual animality is to be identified with a lower part of the soul, or our existence in the phenomenal realm. However, even though dual-process models in psychology might be seen as a modern version of the ancient old parts-of-the-soul doctrine (cf. Haidt 2006), it seems to fair to say that the parts-of-the-soul theory does not find much supporters anymore today, both because of philosophical arguments against such dualisms (e.g. Heidegger 1927), and because of scientific findings. For example, the influence of unconscious processes, moods, emotions, habits, and energy clearly does not stop half-way, but is rather pervasive in all our thinking and acting. Moreover, if there were such things as autonomous souls, they should manifest themselves at least in the morally crucial situations of Milgram’s obedience experiment and Zimbardo’s Stanford prison

---

<sup>5</sup> Of course, Aristotle is a weak naturalist according to some interpretations. But this is not important for the point I want to make: My point is that human nature has often been conceived of in a dualistic way in the history of philosophy, and Aristotle is one proponent of such a dualistic conception.

experiment – but apparently, the behavior of at least most of the participants is guided by situational factors.

So I insist that the problem I described is really new. Or, more precisely, the problem is as old as our self-conception as responsible autonomous individuals, but it has become visible only recently. I suspect that there are at least two reasons why the problem has only recently become manifest: One reason is that the tension has been hidden behind the dualism of soul/mind vs. body, so that all phenomena were just interpreted as falling either in the category of soul/mind or body. However, nowadays such dualisms are widely rejected, and so it becomes a problem that we continue to conceive of ourselves as autonomous moral agents. And another reason why the problem has only become visible recently is the more or less recent emergence of controlled experiments and of statistics in the cognitive and social sciences, supporting the view that we are habitual animals influenced by unconscious and environmental factors. For, clearly, the crucial influence of unconscious factors could not have been seen by a self-conscious reflection or a reconstruction of one’s intuitions.<sup>6</sup>

Now once it is acknowledged that there is an important problem, the next question is how to solve it. There are several options (cf. Alfano 2013). The first option would be to embrace what might be called “dogmatism”, that is, the view that our classical self-conception needs to be defended by all means (cf. Mele 2009). But this would not be a promising way to go, I contend, because it neglects a whole dimension of human existence and would refuse to make the progress scientific results allow us to make. The second option would be a view which might be called “scientific skepticism”, that is, the view that our classical self-

conception has to be completely replaced by a scientifically informed account (cf. Churchland & Churchland 1999). However, this option is not promising either, I contend, because our classical self-conception seems to be indispensable for living the lives we live. Consequently, I propose taking a third path: Using the scientific results in order to question what it means to be an autonomous moral individual in the first place, and asking how autonomous selfhood is constituted. As I will suggest, enactivism and Mead’s pragmatism provide exactly the right resources for this developing a fruitful answer to this question.

But before turning to this important task, it will be helpful to quickly discuss and refute a view I suggest calling “harmonism” – a view which denies the existence of the problem I just discussed ...

### 3. Against Harmonism

There is no one who calls herself a “harmonist”, but I contend that this is an apt name for a view which some philosophers adopt implicitly. Harmonism would be the view that we human beings are just natural parts of a non-dualistic natural world, and that autonomy, free will, reason, agency, and selfhood can unproblematically be conceived of as parts of this natural world. According to this view, humans naturally have eyes and can use them for seeing, and have lungs and can use them for breathing; and in the same way, harmonism assumes, humans have brains and can use them for moral reflection, and have free will and can use it for making moral decisions. Consequently, the problem I described would not really exist.<sup>7</sup>

<sup>6</sup> Obviously, the respective dualism of soul/mind vs. body is already crumbling for over a century (one might think of Nietzsche’s criticism), and controlled psychological experiments also already exist since some time – so it might be not surprising that versions of the new problem already showed up during the 20<sup>th</sup> century, for instance in Freud’s and Lacan’s works.

<sup>7</sup> Of course, it is logically possible to subsume a lot of very different philosophical positions under the concept of harmonism. So, one might find the concept unhelpful for categorizing philosophical views. But I do not claim that my distinction between harmonism and disharmonism is the exclusive, one and for all times best way for making such categorizations. Still, the concept of harmonism can be very helpful for making explicit a widespread, momentous, and non-trivial way of thinking which is often implicit and which is sometimes even

Indeed, harmonism can be arrived at by a misunderstanding of anti-dualism, so that even some pragmatists might get misled into harmonism. I agree with most pragmatists, phenomenologists, scientists, and others that dualism, and in particular the soul/mind vs. body dualism, or mind vs. matter dualism, is deeply problematic. But still, I submit, it is a matter of fact that we conceptualize ourselves as autonomous individuals, and that this self-conception is in conflict with our biological and social nature as it is described by the cognitive and social sciences. But this does not presuppose any problematic metaphysical dualism: we can just accept that any plausible theory of human nature has to conceive of human nature as non-dualistic, while we must also acknowledge the hard fact that our self-conception is in tension with a more empirically informed conception of human nature. Long ago we have construed a conception of a flat earth in the center of the universe, and now we know that this was wrong, and we have given up that conception. But long ago we have developed the conception of ourselves as responsible moral selves, and now this conception runs up against scientific findings. Yet we have not given up that self-conception, and so crucial conflicts arise. Thus, I contend, anti-dualism is right, but still there are important disharmonies within our self-conceptions, so that harmonism is wrong.

Let me briefly explain in a bit more detail why harmonism is so mistaken. One version of harmonism has just recently (even though implicitly) been embraced by a thinker with great merits in pragmatism and in the philosophy of cognitive science, namely by Mark Johnson. In his book *Morality for Humans* (Johnson 2014), he uses findings from the cognitive sciences to argue against a Christian understanding of morality. Using Dewey’s famous metaphor, Johnson insists that we humans are no “little gods” who would have epistemic access to absolute, transcendental values. Johnson insists on, and provides empirical evidence for

---

invisible to those who hold on to it.

statements such as “There is no such thing as a faculty of pure practical reason” (Johnson 2014, 24), “Faculty psychology is outdated and misleading”, (Johnson 2014, 24), and “Moral absolutism is profoundly mistaken” (Johnson 2014, 26). Caricaturing Johnson’s position a bit, he thinks that such problematic ideas can just be cut out of our moral self-understanding like a tumor, leaving back a healthy patient. Once we get rid of the belief in transcendental values and similar things, we can realize how our moral values actually emerge from our very mundane bodily interactions, how moral reflection is just a version of everyday imaginative problem-solving, and so on. So our self-understanding as moral agents would (once it is cleansed of Christianity) be in a perfect natural harmony with our nature as bodily beings.

Even though I strongly agree with Johnson’s project of approaching morality from the point of view of cognitive science, and even though I agree with his rejection of faculty psychology and transcendental values, I am deeply skeptical about the harmonist picture Johnson draws. First of all, harmonism naïvely take it for granted that humans can be agents, or persons, or selves, with the ability to think. Harmonism just takes for granted that humans are autonomous individuals, assuming that morality is just some later addition. In this way harmonism places itself, of course, in the early modern tradition and follows the lead of Thomas Hobbes (1651), for instance. But given the problematic nature of our status as autonomous individuals described earlier, one cannot just assume the existence of autonomous individuals which is called into question elsewhere: consequently, it has to be explained how human animals can be autonomous individuals. This first problem leads to a second problem, namely that harmonism seems to be historically naïve. Friedrich Nietzsche (1886), for instance, has provided an account of what one today would call “social mind-shaping”, an account of how social, and in particular religious practices have shaped and transformed our human abilities, maybe creating even new capacities such as conscience. Then, the harmonist’s idea that Christianity

can be just cut out from our moral practices would be naïve – it would have been the historic condition of the possibility of some of our moral capacities, and it would still be prevalent in our embodied habits for engaging in moral practices. Thirdly, harmonism is metaphysically naïve: clearly, a Cartesian metaphysical dualism of a *res cogitans* and a *res extensa* is metaphysically dubious. But a Platonistic conception, according to which the whole world is in natural harmony and is perfectly ordered, is metaphysically even more dubious. It must be acknowledged that our self-conceptions and our biological nature can easily be in conflict. And fourthly, harmonism is politically naïve. As Nietzsche and Foucault have pointed to, for instance, there are many dark sides to human nature, with social power relations suppressing many human beings. It might of course be right that Nietzsche and Foucault overplay these points. But likewise, it seems, harmonism underestimates them, painting an all rosy picture of humans as natural-born autonomous individuals, in harmony with themselves and with their natural place in a logically well-ordered world.

Consequently, one should give up harmonism. Anti-dualism is right: there is no mystical dualism in human nature. But likewise, harmonism is wrong: There is no mystical perfect natural harmony in human nature.

So we are back to where we were at the end of the last section: We find ourselves faced with a conflict between two important views of human nature, and we need a new account of moral selfhood to solve that conflict. In the next section, I will introduce enactivism as the first building block for a new understanding of moral selfhood.

#### 4. Enactivist Cognitive Science

This section introduces enactivist cognitive science as a first building block for construing a new account of moral selfhood, an account which is compatible with the findings from the cognitive and social sciences, but which avoids the naïve harmonist’s assumption that

humans are just autonomous agents. Enactive cognitive science is part of “4e cognitive science”, which suggests viewing cognition as embodied, embedded, enactive, and extended. Enactivism is influenced by work in radical constructivism, hermeneutics, phenomenology, Buddhism, and neurobiology, and has been introduced by Francisco Varela in his book *The Embodied Mind*, co-authored with Evan Thompson and Eleanor Rosch (Varela, Thompson & Rosch 1991). Enactivism shares pragmatism’s rejections of dualism, of representationalism, and of cognitivism. It also opposes the idea that human behavior is essentially guided by a brain which computes contentful representations; rather, the idea is that there are no contentful states in the brain, and that the brain is not the organ for thinking, just as the heart is not the organ for emotions (of course, there are no emotions without the heart and no thoughts without the brain). In a positive vein, there are at least two constructive ideas which are crucial for enactivism. One idea is that action is essential. But here, “action” is not to be understood as morally autonomous piece of behavior, carried out by a self-conscious human person with deliberate intentions and in the light of moral reasons. Rather, “action” here refers to bodily interactions with one’s environment, interactions even simple animals such as bacteria are capable of. For instance, it has been suggested that at least most human behavior consists in such more or less simple bodily interactions (Hutto & Myin 2013). It has been suggested that action and perception shape each other, for instance that we perceive our environments in terms of action possibilities (Gibson 1979, Noë 2004, 2012). And it has been suggested that being in interaction with one’s environment actually gives rise to the conscious experience we have (O’Regan & Noë 2001, O’Regan 2011). Today there are many different versions of enactivism, but this first idea is widely shared. By contrast, the second idea is less widely accepted, even though it is the idea which gave enactivism its name, and which is a key idea in Varela’s original work: This is the idea that experience is always enacted, that is, made, or

brought about (Varela 1988, Varela et al. 1991). The idea is that it is wrong to suppose that we can truly or falsely represent within our minds a mind-independent, objective reality, which would continue to exist in the same way if the experiencing subjects were removed (so that Cartesian skepticism becomes a possibility). Rather, we only have the experiences we have because we are actively engaging our (mind-independent) environments against the background of subjective background conditions, and thus bring the (mind-shaped) experience of reality about. For instance, when perceiving a purple flower, we might naively think that the purpleness is just there in the outer world, and that we might represent it truly or falsely in our inner world. But enactivism would emphasize that the color experience is actually something we make, for instance because we have human eyes, have the color concepts of our culture, and are able to move around (Varela et al. 1991). By contrast, a person from another culture, or an animal from another species, might experience light of the same wave length as a different color. According to this second idea of enactivism, “making” experience, or making sense, is something which occurs all the time in any living being’s life (cf. Thompson 2007, 2014).

In the following paragraphs, I will present what one might call “an enactivist account of basic agency”. I will concentrate on basic animal action and a conception of an “enacting animal self” because this is what is most helpful for developing a new account of moral selfhood which is in line with cognitive and social science. I will combine enactivism with some ideas from ecological psychology and from the so-called sensorimotor theory of consciousness, and I will take the liberty to add a few new ideas.

A helpful start for developing an enactivist account of basic agency is to be found in Hutto and Myin’s enactivist manifesto, where they write, as already quoted above: “the vast sea of what humans do and experience is best understood by appealing to dynamically unfolding, situated embodied interactions and engagements with worldly offerings” (Hutto & Myin

2013: ix). But whereas Hutto and Myin proceed by arguing against traditional theories, my focus here is more constructive. If one wants to develop a more elaborate enactivist account of embodied agency, one has to ask what “worldly offerings” are, how the dynamic “interactions” with them are to be understood, and how the embodied human agents are to be conceived of, if they are not Cartesian cognizers.

Drawing on work from ecological psychology (Gibson 1979), which is in turn influenced by Gestalt psychology and phenomenology (Merleau-Ponty 1942), I first propose a way to better understand the “worldly offerings”. The key idea, which is shared by many authors from the traditions just mentioned, is that environments we humans find ourselves situated in when we are absorbed in our everyday routines do not consist of abstract objects like a 50 cm high, hard black block. Rather, we immediately perceive our environments in terms of possibilities of, and even demands for actions, seeing a chair as inviting sitting, or a glass of water as inviting drinking. While talking, one might get thirstier and thirstier, and the glass nearby might become a strong invitation to drink from it. According to Gestalt psychologist Kurt Koffka (1935, 7), “[e]ach thing says what it is [...] a fruit says ‘Eat me’; water says ‘Drink me’; thunder says ‘Fear me’.” Less metaphorically, and following phenomenology (and recent cognitive neuroscience) one might call the action invitations experienced by embodied agents in interaction “solicitations”. A glass of water, for instance, can solicit drinking from it. Deliberately departing from the established way of using the term, I suggest calling potential solicitations “affordances”. For example, a fruit, which would be a solicitation for me if I were hungry, is no solicitation but only an affordance for me when I am not hungry.

Some thinkers working in the tradition of ecological psychology hold on to a strong realism about affordances, assuming that affordances exist in a mind- or creature-independent way (Gibson 1979, Noë 2004, 2012). Other thinkers emphasize the importance of

animals' abilities for acting on particular affordances (Turvey & Shaw 1979, Rietveld & Kiverstein 2014). Now I suggest going one step further and combining the ecological theory of solicitations and affordances with the second, more ambitious idea of enactivism outlined above. Then, affordances would not exist mind-independently. Rather, they have to be understood as environmental relatum of an animal-environment-relation. There are, I suggest, subjective background conditions of a particular animal which constitute that that particular animal can enact particular objective environmental features as solicitations. By analogy, a human animal (arguably) has to have the concept "purple" in order to enact light of a certain wave length as looking purple. Likewise, a human animal has to be able to digest milk, and be able to grasp a glass, in order to experience a glass of milk as a solicitation to drink from it. If this is so, it seems that there are many such factors that contribute to which objective situational factors are solicitations for a particular animal, for instance the animal's abilities, habits, expectations, moods, energy, and so on. To have a name for it, I suggest calling the subjective background conditions for enacting objective environmental factors as solicitations "acceptances". For example, a book on quantum physics is only an affordance for a given agent if she understands the basics of quantum physics – her ability of understanding quantum physics would be an acceptance in virtue of which the book would be an affordance in her world. A performance of an opera can be an affordance to attend it for an agent who has the habit of occasionally attending opera performances – her habit would be an acceptance in virtue of which the performance is an affordance in her world. "Affordances" and "acceptances" are then the names of the relata of inseparable affordance-acceptance relations. If this is accepted, at least for the sake of argument, it would – I suggest – allow for developing an anti-cognitivist conception of an embodied agent, or of an acting, active self. The active subject of embodied interactions would not be a ready-made, natural born

autonomous individual. Rather, the subject of embodied interactions would be a mere animal, a bodily organism, to be characterized by the entirety of its acceptances. In this way, a one cell organism could be an embodied agent when it swims towards nutrients, and its subjective contribution to the environmental interaction would consist in the facts that it enacts the nutrients as solicitations against the background of its acceptances, and that it reacts to the solicitation to swim towards the nutrients by swimming toward them.<sup>8</sup> We can very well say that the self of embodied interactions is an enacting animal self, constituting a whole world of affordances against the background of its acceptances, acceptances which make it the self it is.

This notion of an enacting animal self is different from, but compatible with the phenomenological notion of embodied subjectivity. Drawing on work by Husserl and Sartre, Dan Zahavi (2005, 2014) has suggested that all our experiences and interactions carry a minimal trace of mineness, a non-thematic feeling that it is me acting. Certainly, we need such a feeling in order to act successfully, because we must be able to distinguish which changes in perception are the result of a changing environment, and which are the results of our own movements (Vosgerau & Newen 2007). But Zahavi's notion of embodied subjectivity is not concerned with the agent, the active element in minimal embodied interactions, and this is why the new notion of an enacting animal self is (I suggest) a very helpful addition.

Yet it is important to note that the conception of the enacting animal self is importantly different from other conceptions of selfhood: The enacting animal self does not have the reflexivity which normally associated with the concept "self" – the enacting animal self is not self-conscious at all. To explain: When an animal interacts with solicitations, it might of course have a minimal experience of mineness, a sense that it is it who is acting – that is, that it is an embodied self in the sense

<sup>8</sup>This proposal can also be understood as contributing to the important research on "primitive agency", see Frankfurt 1978 and Burge 2009.

described by Zahavi. But the animal does not know that it only experiences solicitations because it has enacted them. The animal only experiences invitations for action, but has no sense for what makes experiencing those invitations possible in the first place – the world-constituting power of its acceptances.

This conception of an enacting animal self is based on an important distinction which is not made explicitly by enactivists, but which can be viewed as rooted in Varela’s thinking: the distinction between the enacted contents of our experiences, and the mechanisms for enacting the contents. Departing from “radical” enactivists who, at least for “basic minds”, argue against using the notion of content in the form of the truth conditions of propositions, I suggest that it is helpful to analyze all experiences as a having (often non-propositional) minimal content. “Minimal content” is just meant to conceptualize the phenomenon that the world shows up as meaningful for us. For example, a solicitation can be the minimal content of an experience (cf. McDowell 2007, 2009). But this content would be different from the conditions of the possibility of the content, from the mechanisms for enacting the content. For example, we might experience a purple flower, but among the mechanisms for enacting this experience is that we possess the concept “purple”, something we do not experience.

Turning to the nature of action, I suggest that action, as a biological phenomenon, consists in a dynamical interaction with solicitations. Action would be a constant sensorimotor loop of enacting some objective features of the current situation as solicitations, of reacting to them by doing what they solicit, of enacting new solicitations, and so on. For example, a dog might spot a squirrel and perceive it as a solicitation to chase it, and react to this solicitation by chasing it; on the way, the dog might spot another dog and perceive it as a solicitation to get in contact, and react by getting in contact, and so on.<sup>9</sup>

---

<sup>9</sup> Of course, fully developing such a sensorimotor

In order to be able to perceive and interact with solicitations, we have to have an implicit knowledge of so-called sensorimotor contingencies. Or this is at least the suggestion of the sensorimotor theory of consciousness which I propose to integrate at his point (O’Regan & Noë 2001, O’Regan 2011). A sensorimotor contingency is the law of how our motor movements relate to change in sensory input (O’Regan 2011). For example, the sensorimotor contingencies are different depending on whether one sees, hears, or tastes something, and on whether one hears a symphony or a bird. It seems obvious that we have to master such contingencies in order to cope successfully with the physical world.

The outline of the enactivist view on basic bodily interactions might hold true even of a lonely bacterium which is the last of its kind. But it is important to note that there is sociality even at this level. As Shaun Gallagher (2001) has emphasized, there is the important phenomenon of primary intersubjectivity – for instance in affective bodily interactions of baby and mother. According to Gallagher (2009), primary intersubjectivity starts at birth, entails sensorimotor capacities for child–other interactions, and is found in neonate imitation, infant’s interactive eye-tracking and discerning expressions of emotions.

This enactivist account of embodied interactions is, I propose, the first building block for a new approach to moral selfhood which is in line with cognitive and social science. It should be clear at this point that the enactivist account I proposed is not only compatible with the scientific findings mentioned at the beginning of this paper – the enactivist account even offers a conceptual framework for systematizing and integrating them. For example, the situational factors which have been found to influence behavior tremendously can now be conceptualized as affordances and solicitations. The many different unconscious “inner” factors which have

---

account of action, and showing its advantages, is an ambitious task, which cannot be accomplished here (but see Weichold 2015: ch. 3).

been found to influence human behavior can be conceptualized as acceptances. The non-existence of a real autonomous disembodied individual self behind actions is no surprise for the proposed version of enactivism, because all what is needed for bodily interactions is an enacting animal self. Finally, the fact that most reason-giving is just a post hoc rationalization can be easily explained by pointing out that most interactions just consist in reactions to solicitations, and not in carefully carrying out plans resulting from conscious deliberation about moral reasons.

However, it is still essential for our everyday life that we do conceptualize ourselves as autonomous individuals who think about moral reasons, are responsible for their deeds, and can autonomously criticize the social order. This was part of our initial problem. And so far, it seems far from clear how enactivism should solve that problem, given that it does not seem able to make sense of our self-conception as autonomous individuals. What we need, I propose, is a further building block for a new account of moral selfhood – and this further building block is Mead’s pragmatist account of social self-constitution.

### 5. Mead’s Theory of Social Self-Constitution

Mead is introducing an important new explanatory resource, namely society. But Mead is going further than others: Mead is not only interested in how two or more humans can make sense of their environment together (“participatory sense making”) (de Jaegher & di Paolo 2007), or how humans understand others (“social cognition”) (de Bruin & de Haan 2012), or how humans can intentionally cooperate as a group (“collective intentionality”) (Searle 1990, Gilbert 1990). Instead, Mead is going much deeper by investigating how social structures constitute the individual self, and how social structures constitute individual action. According to Mead, there are no ready-made, natural-born autonomous individual selves who could later team up in order to construe a society – rather, it is the other way

around: It is social structures which transform human animals into autonomous self-conscious individuals (Mead 1934, 222ff.).

In this section, I will provide an own interpretation of Mead’s account of social self-constitution. Most of what I say will be familiar and uncontroversial. Yet I will add a few new ideas and interpret Mead through enactivist lens. This, at least, will prove very helpful for construing a new account of moral selfhood. Moreover, it will be helpful for bringing out the innovativeness of Mead’s position, which is easily missed, partly for systematic reasons. This is because it is difficult to correctly interpret Mead’s claim that the self is constituted by society. In one interpretation, this would be correct but trivial: Surely, it depends on society that we are members of a family or citizens of a state. If this were so, however, society would only provide some social clothing for pre-existing autonomous agents. And this would not be an interesting hypothesis. In another interpretation, Mead’s account would be false: If individuals are in their very being constituted by the social roles they play, we must ask who is doing the role playing, and who is engaging in social relationships. So claiming that society constitutes the beings which are necessary for society to exist would be circular. By contrast, I aim to provide an interpretation of social self-constitution which steers a middle way between triviality and falsity, relying on the enactivist account of animal selfhood outlined above.

Mead starts where enactivism stops. Just like enactivism, Mead assumes that humans are essentially animals shaped by evolution, bodily biological beings which are in constant interaction with their environments (Mead 1934). Mead calls human organisms “individuals”, but this should not mislead one to assume that Mead believes in the existence of natural-born autonomous agential moral individuals (Mead 1934, 1). The notion of an enacting active self might arguably be used for having a more concise conception of how Mead views our pre-social existence.

Even the account of action implicit in Mead’s writings (Mead 1934, Mead 1938) bears striking similarities to the enactivist account I have outlined. This is the case partly because Mead holds on to an improved version of behaviorism (Mead 1934, 2). But partly, I contend, this is also the case because Mead was a very careful observer of the phenomena of human existence. In any case, Mead suggested that human behavior essentially consists in responding to environmental stimuli (Mead 1934, 6). But which hypothetical stimuli a particular organism is sensitive to depends on subjective factors, which Mead conceptualizes as impulses and attitudes (Mead 1934, 5f.). Moreover, Mead thinks of behavior as constant interaction of organism and environment. Those ideas are particularly well-developed in Mead’s *The Philosophy of the Act*, where he also put forward an action-oriented theory of perception. For example, discussing the example of seeing a penny, Mead (1938, 128) suggests: “We see the oval penny as round [...]. A movement will give us the round penny, and there is the same entrance of the movement as that which we find, for example, in seeing the hammer as that with which we drive the nail. The percept is a collapsed act.” At least if the matter is viewed from some distance, Mead’s accounts of action and of action-oriented perception are very similar to the enactivist conceptions outlined before. Mead’s stimuli correspond to enactivist’s solicitations, and Mead’s attitudes to the acceptances. If one zooms in, however, more subtle and detailed points of convergence and divergence become visible. But even though an elaborate comparison would be a promising philosophical project in its own right, this paper has to move on to discuss Mead’s account of the emergence of selfhood.

As far as the nature of selfhood is concerned, Mead is most interested in the question how the human animal can become an object to itself (Blumer 1968, Blumer 2004: 57f.). In a way, it can be said that Mead has discovered a new philosophical problem here. According to the philosophical agenda of a Cartesian,

there is just no problem of becoming an object to oneself, of thinking of oneself: A Cartesian thinker has just to turn his focus of attention inwards in order to have privileged epistemic access to its own nature as *res cogitans*. But if human beings are primarily animals whose attention is focused on the solicitations of the environment, and who have no sense for the subjective background conditions which contribute to the constitution of the experienced environment, we must ask how it is that humans can think of themselves as objects. An animal can perceive solicitations to flee, drink, and eat. But a human animal can also entertain the thought that he weighs so-and-so-much kilograms, that he would prefer to weigh 5 kilograms less, and that he therefore should not react to the solicitation of the chocolate cake over there. We are not only placed in an environment, we can explicitly think of our selves as objects placed in an environment; we can have explicit conceptions of who we are. We might think of our selves as friends, teachers, writers, or team-members – but the crucial question is how it is possible to think *of* our selves at all. How can human animals develop such an objective-self concepts? In a way, this might be understood as an empirical question to be answered by developmental psychology. Be that as it may, it is important that Mead offers an unusual and potentially inspiring new way of thinking about this question, a way of thinking which might guide future empirical research.

Mead’s answer is that human animals develop explicit self-conceptions by playfully getting into social structures which transform the human animals’ minds. Or, more precisely, the social structures transform the human animals’ acceptances, and *constitute* the human animals’ mind. We have already seen that there is a bodily primary intersubjectivity. But Mead contends that there is a whole new level of sociality once humans get themselves immersed in the world of language. Approaching the phenomenon of language from bottom up, Mead locates the origin of language in the usage of gestures, and in particular in vocal gestures. Analyzing Mead’s account in detail would be a topic in its own

right, partly because Mead might be seen as offering resources for a new account of the natural origins of linguistic meaning, an account which might be superior to, say, classical teleosemantic accounts. But in this context, it is necessary to make a long story short. For a small dog, a thunderstorm can be (minimally and pre-linguistically) meaningful by being a solicitation for fleeing. Likewise, another, large and dangerously looking dog can be a solicitation for fleeing. But likewise, a large dog which looks ambiguous can become a solicitation for fleeing when the large dog does something, for instance when it snarls. And now, it is even possible that the large dog becomes a solicitation for fleeing when it makes a sound, for example when it barks intensely. But this last case is importantly different from the previous cases. For, the barking has not only a meaning to the small dog, by being a solicitation for fleeing. Rather, the barking also gets a new kind of meaning to the larger dog who hears its own barking: it can now realize that its own barking can make other dogs flee. So this is knowledge of what one can do with sounds – it is the knowledge of a social meaning, of what one can make others do by producing vocal gestures. The meaning of a vocal gesture consists in the social consequences of its usage – an account which brings Mead in the vicinity of Wittgensteinian thinking (cf. Wittgenstein 1953: §43). Then, this is how, very roughly, Mead thinks that more complicated forms of linguistic meaning arise from the more minimal forms of meaning in embodied interactions.

Using Wittgensteinian terminology for expressing a Meadian point, it can be said that human social life consists in playing language games (Wittgenstein 1953: §7). For instance, we play the language games of selling goods, of having a discussion, or of giving a talk. In these language games, the rules of the games define what moves in the games we can make: For instance, we can ask for a price, make an offer, buy a product, and so on. Moreover, different human animals have different functions and different social roles in these games, which define which moves in the game they are allowed

to make: The customer is not allowed to just carry away the goods she likes, and the shopkeeper is allowed to fix the price. In addition, there are often what one might call “materialized behavior settings”: a shop is the place for the language game of selling, a room in the university the setting for academic discussions.

If this is on the right track, it would mean that we have to distinguish between two notions of action. On the one hand, there is embodied agency, the embodied dynamic interactions with a material environment which have been described before, and of which even primitive organisms are capable. On the other hand, there is social agency, the making of moves in language games, which requires the embeddedness in social practices. If this is true, even individual actions, such as buying a good or preparing a talk, are social affairs (cf. Gergen 2009).

As it has been suggested before, organisms learn to successfully perform embodied interactions by exploring their environments and by seeing what happens in response to their bodily movements, so that the organism eventually grasp the relevant sensorimotor contingencies. Now how do young human children who already master relevant sensorimotor contingencies go on to eventually make moves in language games? Do their brains grow capacities like “the will” and “reason”, so that the children first become full-fledged agents, who then, in a second step, use their agency to engage in language games? But we have already seen that such a view does not find much support from the sciences, and we have agreed with Mark Johnson that such a presupposition of faculties is outdated. So I want to suggest an arguably more plausible explanation. The idea is that human animals not only have to learn mastering sensorimotor contingencies, but also to learn mastering sociocultural contingencies. Mastering a sociocultural contingency would mean to know which bodily movements count as movements in a language game, and which social consequences making those moves in the language game have (Weichold 2016). Using Mead’s account of the meaning of vocal gestures, we can say that this is how a child understands the

meaning of the moves in language games. Obviously, learning to make moves in language games depends on social feedback, on what others acknowledge as counting as a move in a language game. And it suggests itself that young children learn to master such sociocultural contingencies by exploring their social environments playfully. Clearly, this “mastery” or “knowledge” of sociocultural contingencies is a bodily knowledge (like “knowing how” in an anti-intellectualist interpretation, cf. Dreyfus 2005, Noë 2005), and not a grasping of contentful true propositions (*pace* Hutto & Myin 2013, 24-32).

Consider a baby which already masters sensorimotor contingencies. It produces weird sounds which have no particular meaning to anyone. But then, by trial and imitation, the baby produces the sound “mom”. The production of this particular sound suddenly has social consequences, and the baby suddenly receives attention and care in reaction. In this way, the sound acquires a meaning for the baby. It is a social meaning, depending on the social reactions of others.<sup>10</sup> In a similar way, we learn to play new language games throughout our whole lives. For instance, first year students at the university often do not know which kind of help they can ask for from their professors. So they have to try it out, and learn to master the academic sociocultural contingencies in virtue of being ignored, rejected, or helped, and being taken care of.

It is important that learning to play a language game necessarily goes together with a sense of one’s own position in the game. We have certain positions, or functions, in the language games, and these positions define which moves in the game we should and are allowed to make. Those positions are our roles. Even though the idea of using theatrical metaphors for analyzing human life is already found in antiquity, and

even though the notion of a social role is already discussed in Confucianism and British Idealism (Bradley 1876), Mead has to be credited for making this important concept prominent in philosophy and social theory. Most of the time, we play roles such as being a friend, daughter, teacher, or sports team member. For example, when we are in a shop, we might play the role of a customer, and in this role, we can make the moves of buying a product, or making an inquiry about a good. But singing a song, just eating a product, talking to the shopkeeper about the meaning of life, and many other things are not legitimate moves in the particular language game.

Importantly, which precise moves in a game we can make depends on which moves others allow us to make, on how they see us. One can only grade papers if there are others who accept the grading as legitimate, who see one as occupying the role of a teacher. And what exactly others allow us to do also depends on how interesting, smart, rich, or noble-born others take us to be. Consider one and the same new philosophical idea which is once proposed in a talk by a famous American professor, and once by an undergraduate student. It is not unlikely that the professor is praised for brilliance in exploring new ways, whereas the undergraduate will be blamed for departing from orthodoxy. So the American professor will think of herself as a genius afterwards, while the undergraduate might think of herself as a mere beginner.<sup>11</sup> Then, this should show: Who we are depends on who others take us to be.<sup>12</sup> Our positions in the games of life depend on which positions others take us to have, which positions we think others view us as having, and which positions we implicitly accept. In a way, this is a boiled down, social enactivist version of Hegel’s master-slave-dialectic. And so it might now be

---

<sup>10</sup> Based on work from Michael Tomasello (2016), one might say that *human* animals possess a *higher degree of shared attention* than other animals – and this higher openness for social learning is what enabled humans to develop an unparalleled world of culture.

---

<sup>11</sup> Let us assume for the sake of the argument that the American professor and the undergraduate student show equal skills in defending the idea against objections.

<sup>12</sup> Even the following, more complicated version is true: Who we take our selves to be depends on who we take others to take our selves to be.

tempting to say that we can only recognize our selves in the eye of the other. But this would be too weak, because others *constitutes* our selves in the first place, constitute that we exist in the social world at all.

This is why Mead suggests that we have to take the role of the other in order to have an objectified conception of ourselves (Mead 1934). I have to put myself mentally in the shoes of the shopkeeper to see myself as a customer, and to put myself in the shoes of a student to see myself as a university teacher. However, often we have already internalized the attitudes of the others, and then it might be enough to just hear one’s own speech: When I hear myself having to beg for a small favor, I understand that my position in social space is inferior to the position of the other. When I hear myself giving instructions, I understand that I have a higher position.

Mead suggests that children develop such an objectified conception of their selves by playfully engaging in role play and by taking the attitude of others towards themselves. In the first stage, the “play” stage, children just playfully explore different roles. But later, children move on to a second stage, the “game” stage, where they grasp all the relevant sociocultural contingencies of a particular game. In reference to children, Mead says: “In that early stage he passes from one rôle to another just as a whim takes him. But in a game where a number of individuals are involved, then the child taking one rôle must be ready to take the rôle of everyone else” (Mead 1934, 151).

In the further development, children even develop the conception of the so-called “generalized other”: “The organized community or social group which gives to the individual his unity of self may be called ‘the generalized’ other.’ The attitude of the generalized other is the attitude of the whole community” (Mead 1934, 154). The idea is that we understand the rules of the whole game, of the whole social institutions, and can abstract away from particular others. We might then have a sense of being a teacher independently from which particular students show up, or of making an

academic proposal independently from whom exactly is in the audience.

This, then, is how we develop an objective conception of ourselves. We acquire a sense of being someone, of existing in and having a position in social space, through others treating us as being someone. Our sense of being someone solidifies once we master whole games and play more or less fixed roles. And by the time, we can abstract away from particular roles and particular others, and think of our selves as being particular persons who can play this role or that.

If we interpret Mead creatively, this is also, roughly at least, how we acquire our self-conception of being autonomous individuals. We are playing games such as making moral demands, criticizing and changing social norms, and distancing ourselves from certain social roles (cf. Mead 1934, 168). Abstracting away from the details, this shapes our conceptions of being persons: we now think of ourselves as autonomous individuals.

By providing this account, Mead also offers a highly unusual and very innovative conception of self-consciousness. Self-consciousness is no longer understood as the most private thing, something a Cartesian subject knows with first person authority once it turns its gaze of attention inwards. Rather, self-consciousness is a public affair. It is the consciousness of a conception of oneself which one has developed by abstracting away from one’s social positions which, in turn, are constituted by how others treat and see one.<sup>13</sup>

Having an objective conception of oneself, and having objective conceptions in general, open up the possibility of interacting with oneself. In a kind of inner dialogue, a human animal can now, so to speak, talk to herself, and regard her own thoughts as solicitations to

<sup>13</sup> In other words, the consciousness of one’s self-conception is genetically constituted in social interactions, is occasionally socially confirmed or disconfirmed, and is often a consciousness of how one would interact with the generalized other. But still, it is possible to think, at one very moment, (in a public language) a particular thought, the content which no one else will ever find out.

respond to. Obviously, this internalization of external practices may give rise to a rich inner life, and to delayed but more considerate reactions to environmental solicitations. Such an inner interaction is of course representational, but if representations are conceived of as special and extraordinary phenomenon, this is in line with enactivism.

Mead’s account of the self can be regarded as a very helpful contribution to the literature on selfhood. Mead shares with no-self positions the rejection of a Cartesian autonomous self (Metzinger 2003, Albahari 2006). However, Mead would insist that it is still important that we have an objective conception of our selves. Mead’s position is very well compatible with accounts of embodied selfhood (Zahavi 2005, 2014). However, Mead’s position goes much further, and is thus a necessary complement. Moreover, Mead might agree with narrative accounts that we develop a unified outlook on our lives by telling stories about our selves. But this already presupposes that we have an objective conception of our selves. Mead’s account is thus also a necessary complement to narrative accounts, explaining how we can be objects to ourselves.

But even though Mead’s account social self-constitution seems to be so promising and powerful, it still cannot provide the whole story about moral selfhood. To see that there is a very important problem still waiting to be solved, I will discuss one last issue, namely Mead’s prominent distinction between “I” and “me”. Mead makes this complicated statement: “The “I” reacts to the self which arises through taking the attitudes of others. Through taking those attitudes we have introduced the “me” and we react to it as an “I”” (Mead 1934, 174). Obviously, the term “me” refers to our objectified self-conception which has developed in the light of how we think others see us. But what Mead exactly means with the “I” has created much puzzlement in the literature (cf. Cook 1993, cf. Lewis & Smith 1980). This is so, partly because Mead also emphasizes that we human beings can play our roles in new, creative, surprising, and characteristic ways, and because he

explains this fact with reference to the “I”. So is Mead, under the label “I”, resurrecting the classical autonomous individual? However, this would run against all the findings from the cognitive and social sciences, against enactivism, and against Mead’s own intentions.

Acknowledging that this is a complicated and controversial matter, here is a simple interpretation of what Mead might have meant with the “I” and the “me” (an interpretation sharing central ideas with the interpretations from Hans Joas (1985) and Gary Cook (1993)). On the classical Cartesian account, the self is subject and object at the same time: it can effectuate action in virtue of its mystical agent causal powers, and it can be an object to itself in virtue of its mystical self-transparency. Now Mead has provided a non-mystical, social account of the self as object. This is his account of the “me”. However, the question has not yet been answered how the self can be a subject, how we human beings can be active, how it is possible to act in light of our self-conception. Mead’s notion of the “I” can be considered as an attempt to solve this problem (cf. Joas 1985, cf. Cook 1993).

However, Mead-scholars such as Hans Joas and Gary Cook have emphasized that this view leads to problems: For instance, Cook (1993, 55) emphasizes that Mead holds that the “I” can be immediately experienced – but then, the “I” would not only be the self as subject, but also the self as object. Moreover, Joas (1985, 88f.) emphasizes that, at least at an early point in Mead’s intellectual development, there is a tension between Mead’s idea that the “I” can be immediately experienced, and his idea that the “I” is socially constituted.

Now I wish to suggest that those (apparent) problems can be avoided if my enactivist notion of an enacting animal self is brought into play. After all, the enacting animal self is a notion of the self as a subject. Yet the enacting animal self is not socially constituted (even if it can be partly socially transformed). Moreover, it is not immediately experienced (apart from the minimal sense of mineness discussed above), and is

instead the un-experienced condition of the possibility of experiencing. Against this background, the notion of an enacting animal self can do the job the “I” is assigned to do – namely, having an anti-Cartesian conception of the self as subject –, while it can at the same time avoid the problems Mead’s notion of “I” has in the textually faithful interpretations of Cook and Joas. Consequently, I propose the charitable interpretation of understanding Mead’s “I” as the enacting animal self. The idea is then that human action always consists in interacting with solicitations against the background of acceptances. This can explain how human animals can be active. Human beings are active in virtue of being enacting animal selves.

Moreover, the conception of an enacting animal self might explain how we human animals can play our roles in new, surprising, and characteristic ways, even if there is no such thing as a creative spontaneous autonomous individual. Importantly, interacting with the world shapes our mind, and thus how we look back at the world (Bourdieu 2000); interacting with the world transforms our acceptances. Now we always act against the background of acceptances, and our acceptances have been shaped in long histories of interactions with different material and social environments. Moreover, we bring different biological pre-conditions to the interactions. So, everyone has different acceptances, and thus enacts her world differently. This is, very roughly at least, why everyone plays her role more or less differently, and how surprising and new interactions can arise. Social structures enable us to become persons in the first place, but they do not preclude individuality. And this explanation is very well in line with Mead’s ideas, in particular if the “I” is interpreted as enacting animal self.

Yet even if this is on the right track, the current analysis of moral selfhood is not yet complete. For, enactivism and Mead both suggest accounts of self- and world-constitution. Both suggest that our environments, and our self-conceptions are enacted, made, constructed, brought about. But this runs vehemently

against the realistic spirit which is so important for our everyday self-understanding. In our everyday practices, we do not think that we are autonomous individuals in virtue of being recognized by others. Rather, we think that we just are autonomous agents with a particular personality, in virtue of our inner nature. We think that our environments are just out there, and are not enacted. And we think that our selves are just in here, and are not constituted. So how can this strong mismatch between the enactivist-constructivist spirit and the realism in our everyday practices be explained? Since realism is so important to our everyday self-understanding as moral selves, solving this problem is essential for developing a new conception of moral selfhood.

Thus, one further and final building block is needed, a building block which consists in a more complicated combination of enactivist cognitive science and Mead’s pragmatism ...

## 6. The Game of Selves

According to Mead’s social enactivism, we human beings construe an objectified conception of our selves as distinct autonomous persons by abstracting from the roles others tread us as having in language games. But in our everyday social practices we do not think of our selves as construed, constituted, enacted self-conceptions. Rather, we think that we just *are* autonomous individuals. Our everyday social practices contain strongly realist conceptions of self and world. Consequently, it has to be acknowledged and appreciated that enactivism and Mead’s account of social self-constitution strongly depart from common sense. Any other judgment would mistake the radical nature of enactivism and of Mead’s account.

In this section I aim to complete the new social enactivist account of moral selfhood by showing how a more complicated combination of enactivist cognitive science and Mead’s account is superior to common sense, and by explaining how it can even explain the

emergence of the realism in our common sense self-conception. To do so, I will start with rehearsing three important ideas which have already been alluded to.

The first idea is that Mead emphasizes that children go through the stage of play and the stage of the game in order to learn how to properly play their roles in serious language games. As described earlier, Mead suggests that children first playfully explore what it is to play a role, and then later learn to cope with the rules of organized games, understanding the interconnection of all the roles involved in the game. Today it is widely agreed upon, not only in Vygotskian psychology (Vygotsky 1978), that playing is essential for a healthy development of personality. But most people might have the intuitions that in time, children learn to distinguish funny games from serious reality. But how does this crucial step happen? What exactly do children “realize”? Is this realization a momentary insight? Now, I want to propose a different view: We humans never stop playing. Our social reality is essentially nothing but an assembly of language games. Of course, the language games which make up our lives have existential importance, and might earn us the money we need to survive, or might be essential for achieving the life goals we might have set for ourselves. So we are often playing our roles in the essential language games not with a playful attitude, but rather with seriousness and a sense of existential gravity. Then, I propose, children do not suddenly “realize” a metaphysical reality behind the social games, but are rather trained to eventually play certain language games with a more serious attitude, and to not treat the relevant language games as funny games.<sup>14</sup> This view, after all, would fit very well with Wittgenstein’s insistence on the importance of language games (Wittgenstein 1953), with Friedrich Schiller’s insistence on the importance of play for being a human being (Schiller 1795), with Mead’s account of the

development of children (Mead 1934), and with Johan Huizinga’s famous anthropological thesis that human culture originated from playing games (Huizinga 1938).

The second idea is the importance of the unconscious. This importance has already been mentioned in the description of the empirical findings in the beginning of the paper, and already Mead, being familiar with Freud’s work, acknowledged the existence of an “unconscious self” (Mead 1934, 163). But it is important to appreciate how deep this idea goes: It implies that we humans might have no sense at all – at least not from introspection or a reconstruction of our intuitions – of the biological and social conditions of the possibility of our experience. We experience the world and our self-conceptions in a realistic fashion, but we have no idea about what actually gives rise to that experience. One might thus speak of “the transcendental unconscious”.

The third idea is my enactivist distinction between the enacted content of our experiences and the mechanisms of enactment. For instance, we might have an experience of purple, but no sense that this experience is enacted, amongst other things, because we have certain (human) eyes and because we possess particular culture-relative color concepts.

Now the three ideas can be drawn together. I will start with an analogy. We might view the game of chess from outside: Then, we see how the game has evolved during the centuries. We see that playing the game only works because different players collectively accept the same rules. And we see that certain pieces of wood count as king, queen, or bishop. But we might also view the game from within, so to speak, and then there are no pieces of wood which *count* as kings, but just kings, and no rules which are valid because they are agreed upon, but rather plainly legitimate and illegitimate, and good and bad moves. Now I suggested that our enculturated human life essentially consists in playing language games. Enactivism and in particular Meadian social enactivism explain how our interactions with material and social environments really work, and how

---

<sup>14</sup> Of course, there is work in educational science on how this transition from playing to having a sense of “reality” takes place – see, e.g., Winnicott’s *Playing and Reality* for a psychoanalytical perspective.

we gain our roles in language games. But once we are fully immersed in the language games, we experience our selves and the world *from within the game*.

Roughly, our “view from within” emerges like this: Others treat us as and count us as being autonomous individuals, or moral agents, or persons with a distinctive personality. We thus learn about our position in social space, learn which moves in language games others allow and expect us to make. We then develop an objectified self-conception, abstracting more and more from particular social roles and viewing our selves against the background of the generalized other. We become used to having this self-conception, our acceptances undergo a social transformation, and this shapes what we experience. We *embody* our social roles and experience reality from the point of view of our role-constituted identities. We start experiencing our selves in terms of the roles we play, as friends, teachers, and moral agents. And then, by the time, we start believing that we just are what we experience our selves to be. We reify our self-conceptions. However, as it should be clear, this way of thinking completely misses the real enacting conditions of our experiences: the acceptances, the sensorimotor and sociocultural contingencies and the constant social interactions. We might be experts in knowing how to master sensorimotor and sociocultural contingencies, but we have no sense for our own mastery. Naturally, we are unconscious of the unconscious background conditions for enacting the experience of reality. So we end up believing just in what we see, and not in what makes this seeing possible. We then believe in an objective mind-independent reality with objective, eternal moral laws, and we take our selves to be autonomous individuals with libertarian free will, agent causation, autonomy, reason, and an inner moral compass. We start living in a kingdom of ends, in a noumenal world – but this noumenal world can now be recognized as a game world.

In one way, this self and worldview is an illusion, and in another way, it creates a new reality. It is an illusion from the point of view of philosophy and the cognitive

and social sciences. After all, we are not really autonomous individuals, given both the evidence from our existential experiences and from the controlled experiments from the sciences. For instance, in our everyday practices we might believe in agent causation and think that we can make decisions without any causal preconditions. But this would be plainly wrong (cf. Pereboom 2007). Of course, philosophy and the sciences are also social practices which formulate their theories from particular points of view. But still, those views are more informed than our everyday self-conception; they take into account controlled scientific findings and our human experiences, and are developed by careful argumentation. Moreover, the content of the philosophical-scientific view is more stable than our everyday self-conception: It holds for all animals at all times that their actions consist in dynamic interactions with solicitations. By contrast, our everyday self-conception is radically contingent in the sense that it would be non-existent if human animals had not started to develop culture in the last few hundred thousand years, and that it would be entirely different if the cultural evolution had been different in the last 3,000 years (or even in the last centuries, insofar as rather new ideas such as individual autonomy are concerned).

But at the same time, our self-conception as autonomous individuals is – I contend – not a mere epiphenomenon. Of course, the self-conception (so to speak) pretends that it is the whole story about moral selfhood and human agency; it pretends that action is to be explained solely with reference to autonomous individuals. This is wrong: What really goes on is to be best explained in terms of embodied interactions with material and social solicitations. But still, the belief that we are autonomous individuals is causally efficacious and plays a role in those interactions. The explicit conceptions others have of us really influence how they treat us, and the explicit conceptions we have of our selves really influence how we interpret our roles. In this sense, we are the beings we are because we play our selves.

There are many ways in which our real-life interactions are influenced by our self-conceptions as autonomous agents with distinctive personalities. For instance, if we have a strong sense of being a creative and independent person, we might interpret our roles, as speaker at a conference for example, differently, and propose more independent ideas. If we believe in objective moral laws and in our being on the right side, we might be more willing to criticize others. If we believe that we have in-born moral rights, we might find ourselves ill-treated in certain interactions. So even if we are no autonomous individuals in reality, believing that we are can bring about much criticism and conflict, which can be a productive element for social change. The other way around, interactions which go well, and being particularly well-treated, might let us develop a high sense of our selves. A famous rock singer, for instance, might reify the enthusiast reactions of his fans into believing that he, as a matter of fact, is a particularly worthy and extraordinary person.

Against this background, it becomes visible that the enactivist median account developed here is very much in line with Mead’s position, but goes beyond it. Like Mead, the account suggested here holds that we human beings are socially shaped animals, which are likewise capable of reflective thinking and self-control. However, Mead does not appreciate that we conceptualize ourselves as pre-social, autonomous individuals with free will in our everyday practices, and that this is contradicted by findings from the cognitive and social sciences (which largely did not yet exist at Mead’s time). Consequently, the critical element of the present proposal – that our self-conception in the game of selves is in one sense an illusion – is missing in Mead’s account. Enactivist cognitive science and Heideggerian and Bourdieuan theories might thus be better resources for developing this aspect.

To sum up: We are autonomous individuals only in the game of selves, just like certain pieces of wood are kings only in the game of chess. We believe that we are pre-social autonomous individuals. But we have this

belief only due to social interactions, by going through the stages of role play and the game, by learning how others see and react to us. The self is born in social space, but it denies its origin.

This social enactivist proposal then connects up to the received literature in interesting ways. Firstly, it allows steering a middle way between eliminativism (Churchland & Churchland 1999) and realism about our folk-psychological self-conception (Fodor 1983).<sup>15</sup> Taking ourselves to be autonomous agents is an important game which enables many important social institutions such as holding people responsible.<sup>16</sup> And secondly, the proposal allows steering a middle way between classical theories which hold that human beings are autonomous persons, and post-structuralism theories which take the autonomous subject to be “dead” and think that social interactions are to be explained in terms of social power. The proposed account suggests that we are no autonomous agents in reality, but that playing autonomous agents is an important game which really makes a difference to human life.

In this sense, the proposal finally offers a new account of moral selfhood, an account which is in line with the empirical findings from the cognitive and social sciences discussed at the outset. We human beings are at root human animals, but we become more by social role play, by receiving a position and a self-conception in the game of selves.

Importantly, this new understanding of human beings as role-playing animals does not leave everything as it is. By contrast, it can sharpen our eye for the many disharmonies, tensions, frictions and conflicts in human nature – disharmonies which are of tremendous existential and ethical importance. In the next and last

---

<sup>15</sup> Of course, Fodor is a realist about representations, content, and mental states, but is just not interested in topics such as selfhood and autonomy.

<sup>16</sup> In this sense, my proposal is also different from Daniel Dennett’s “intentional stance” view (Dennett 1991): According to my proposal, we do not just see certain events as intentional actions from an intentional stance, as Dennett would have it. Rather, we also *play* the roles of intentional agents.

section, I will briefly point to some of those crucial disharmonies, in order to give a sense of the important implications of the new social enactivist account of moral selfhood.

### 7. Unity and Disharmony in Human Nature

This paper started with showing how two views concerning human nature are both important for us, but in conflict with each other. Meadian social enactivism has shown a path for reconciliation: We can think of human beings as human animals who play the roles of autonomous persons in the game of selves. But even though this allows for conceiving of human nature as unitary, this proposal makes intelligible and visible that our human existence is by no means harmonious. Referring back to the earlier discussion about harmonism, the social enactivist account proposed here is both anti-dualist and anti-harmonist. In short, it is perfectly compatible with each other that humans are animals and that humans play games (like many other animals). But our human animal nature is often in disharmony with the *content* of our self-conceptions within the game of selves. In the following paragraphs, I will first discuss the unity of human nature and then point to the disharmony in human existence.

I started developing the Meadian social enactivist account neither from a scientific reductionist analysis in terms of neurons, nor from a naively realist analysis in terms of autonomous persons. Rather, as enactivists and phenomenologists would have it, I started from an analysis of human beings as biological organisms. Enactivism and Mead's account of social self-constitution have shown that there is a natural biological and cultural way from understanding humans as biological organisms to understanding humans as role players with objectified self-conceptions. The social role play transforms our acceptances in part, so that we are essentially enculturated beings. There are neither different parts of the soul (as Plato thinks), nor different substances of human nature (as Descartes would have it) nor different

layers (as Dreyfus (2005) and Husserl (1989) would have it). Thus, according to the account on offer, human nature is unitary.

But this unity in human nature very well compatible with the possibility of disharmony. For, the content of the game of selves can be different from how things really are. Our self-conceptions can be in tension with our biological nature. For example, a person might start thinking of himself as Superman; still, he cannot fly once he tries. But of course, conceptualizing oneself differently from how one is in a biological sense can also sometimes be a good thing and change reality. If we are not born as autonomous individuals with human rights, responsibility, and self-control, it is certainly a tremendous cultural achievement to treat each other as if we were such autonomous individuals (cf. Margalit 1996). For instance, even if we do not have libertarian free will, believing that we do, and that we will be held responsible for our deeds, can really increase our self-control (cf. Nietzsche 1886). In any case, the important point is that the unity of human nature is compatible with disharmonies between our real biological nature and the contents of our self-conceptions.

Indeed, there are many such disharmonies – at least, if the social enactivist point of view suggested here is adopted. For instance, consider the nature of reality. Immersed in the game of selves, we think that there is an objective mind-independent reality which is equivalent to the reality we experience. But from a philosophical point of view (informed by social enactivism), we know that the reality we experience is something enacted, something brought forth against the background of unconscious conditions. Consider the nature of the self. Immersed in the game, we think that we really are the particular autonomous persons we take ourselves to be. But from a philosophical point of view, we are just sense-making animals whose self-conception is constantly reconstructed and modified in the course of social interactions. Consider personal identity. Immersed in the game, we take ourselves to be one and the same person from birth to death. But from a philosophical

point of view, we are ever changing creatures. Consider the body. Immersed in the game, we think that we are persons who happen to be embodied in one particular body, but who could, at least theoretically, be transferred to another body (cf. Parfit 1984, cf. Cassam 2011). But from a philosophical point of view, we are living bodies who happen to have one self-conception, but who could also have another self-conception. Consider sociality. Immersed in the game, we take ourselves to be pre-social autonomous agents. But from a philosophical point of view, we can only play the causally efficacious role of pre-social autonomous agents in virtue of social structures. Consider rationality. Immersed in the game, we think that we have a capacity called “reason” which allows us to tell good from bad. But from a philosophical point of view, considering reasons is imaginatively talking to the generalized other, which is, in turn, an abstraction from the real-life others we happen to have grown up with. Consider moral laws. Immersed in the game, we take moral laws to hold universally. But from a philosophical point of view, what we take to be universal moral laws is an abstraction from the social rules which hold in the communities in which we are thrown. Consider finally moral agency. Immersed in the game, we think that we are, so to speak, only weak-willed shadows of the ideal virtuous agents we could and should be. But from a philosophical point of view, we are human animals who live under the guidance of a view of ideal virtuous agency, which we have socially construed.

Appreciating these disharmonies is of ethical and existential relevance. If we, so to speak, wake up from our immersion in the game of selves, we can recognize the game as a game, and thus get a new sense of who we are as human beings. This can be important in many ways. For example, in a negative vein, we can realize that we often do not have the control over our action which we think we have, when believing in free will or autonomy; situational and unconscious factors always play a crucial role. But in a positive vein, we can also see that we often have much more control over our lives

than we take ourselves to have: We can tremendously influence who we are by changing our self-conceptions, in particular by changing which language game we participate in. Connecting this to ideas from Heidegger (1927) and Sartre (1943), we can say that there is genuine sense of human freedom in being ready to dis-identify with (“to nothing”) the particular role-constituted identity one happens to find oneself thrown into. This way of thinking can even be carried to the social level, and it suggests itself to think: We as a community are collectively responsible for ethically improving and re-designing the social roles which constitute our existence as individual persons. Maybe a better society would not be a society with more moral persons, but a society where the social institution of individual moral personhood is improved. But developing these ideas is, of course, a topic for another occasion (cf. Weichold forthcoming).

## **8. Conclusion**

Summing up, I have started with describing how we human beings conceive of ourselves as autonomous moral selves. This self-understanding is deeply embedded in our everyday moral practices, and it enables important practices such as holding people responsible for their deeds. Yet the self-understanding is seriously challenged by findings from the cognitive and social sciences, as well as by careful observations of human life. Human behavior appeared to be not a matter of agency and autonomy, but of animality and automaticity. Fortunately, enactivist cognitive science and Mead’s pragmatist account of social self-constitution could be teamed up and come to the rescue of our everyday self-understanding. Enactivism was able to synthesize the manifold empirical findings into a unified conceptual framework, and was able to develop an empirically supported view of the self as subject, of human beings as animal agents whose actions consists in a constant dynamical embodied interactions with experienced solicitations in the material environment.

With the help of Mead, it was possible to make intelligible that humans are also embedded in and even constituted by a social environment: Based on the feedback of others, human animals learn which of their bodily movements count as moves in language games, and what those moves mean in terms of their social consequences. This gives humans the idea that they occupy a certain place in the social world, which then leads to the development of an objectified self-conception. Given that humans are so used to thinking of themselves in terms of their self-conceptions, and that this is reinforced by many social interactions, humans reify their self-conceptions and think that they really are what they are taken to be within certain language games. Thus, the Meadian social enactivist account allows thinking of human nature as unitary – but it is a unity in disharmony.

Of course, many important questions are still open, for instance about the details of the social mechanisms in self-constitution, about the potential for integrating more ideas from enactivism and Mead and other sources, about potential applications of the account to topics such as personal identity, free will, agency, rationality, and morality, and about the existential and ethical implications only hinted at.

Yet the social enactivist account can already be considered as bringing with it many advantages: It saves moral selfhood from the cognitive and social scientific challenges, offers a fruitful combination of enactivism, pragmatism, and moral theory, has important existential implications, and opens up a new understanding of what we are as human beings: creatures whose minds are not only embodied, embedded, extended and enactive, but also enculturated, and whose finest cultural products are our selves.<sup>17</sup>

<sup>17</sup> I would like to thank Tyler Q. Sproule, Hannes Worthmann, two anonymous reviewers, and an audience in Paris for very helpful comments on earlier versions of this paper.

## References

- Albahari, Miri. 2006. *Analytical buddhism: The two-tiered illusion of self*. Basingstoke: Palgrave Macmillan.
- Alfano, Mark. 2013. *Character as Moral Fiction*. Cambridge University Press.
- Anscombe, G. Elizabeth M. 1963. *Intention*. 2nd ed. Cambridge, Mass: Harvard University Press.
- Aristoteles. *Nicomachean ethics*. 1999. [NE] Translated by Terence Irwin. 2nd ed. Indianapolis: Hackett.
- Bargh, John A., and Tanya L. Chartrand. 1999. "The unbearable automaticity of being." *American Psychologist* 54, no. 7: 462–479.
- Baumeister, Roy F., Ellen Bratslavsky, Mark Muraven, and Dianne M. Tice. 1998. "Ego depletion: Is the active self a limited resource?" *Journal of personality and social psychology* 74, no. 5 : 1252–1265.
- Blumer, Herbert. 1968/1986. *Symbolic interactionism: Perspective and method*. Berkeley: University of California Press.
- . 2004. *George Herbert Mead and human conduct*. Walnut Creek, California: AltaMira Press.
- Bourdieu, Pierre. 2000. *Pascalian meditations*. Stanford, California: Stanford University Press.
- Bradley, Francis. 1876/2006. H., *Ethical studies*. Oxford: Clarendon Press.
- Bruin, Leon de, and Sanneke de Haan. 2012. "Enactivism and social cognition: In search for the whole story." *Journal of Cognitive Semiotics*, no. 1: 225–250.
- Burge, Tyler. 2009. "Primitive Agency and Natural Norms." *Philosophy and Phenomenological Research* 79, no. 2 : 251–278.
- Cassam, Quassim. 2011. "The embodied self." In *The Oxford handbook of the self*. Edited by Shaun Gallagher, 139–56. Oxford: Oxford University Press.
- Churchland, Paul M., and Patricia S. Churchland. 1999. *On the contrary: Critical essays, 1987-1997*. Cambridge, Mass: MIT Press.
- Cook, Gary A. 1993. *George Herbert Mead: The making of a social pragmatist*. Urbana: University of Illinois Press.
- Darley, John M., and C. D. Batson. 1973. "'From Jerusalem to Jericho': A study of situational and dispositional variables in helping behavior." *Journal of personality and social psychology* 27, no. 1: 100-108.
- Davidson, Donald. 1963. "Actions, reasons, and causes." *The Journal of Philosophy* 60, no. 23: 685–700.
- Dennett, Daniel C. 1991. "Real patterns." *The Journal of Philosophy* 88, no. 1: 27–51.
- . 1992. "The Self as a Center of Narrative Gravity." In *Self and consciousness: Multiple perspectives*. Edited by Frank S. Kessel, Pamela M. Cole and Dale L. Johnson, 103–15. Hillsdale, N.J: L. Erlbaum.

- Dewey, John. 1925/2007. *Human nature and conduct: An introduction to social psychology*. New York, NY: Cosimo.
- Doris, John M. 2002. *Lack of character: Personality and moral behavior*. Cambridge, U.K, New York: Cambridge University Press.
- . 2015. *Talking to our selves: Reflection, ignorance, and agency*. Oxford: Oxford University Press.
- Dreyfus, Hubert L. 2005. “Overcoming the Myth of the Mental: How Philosophers Can Profit from the Phenomenology of Everyday Expertise.” *Proceedings and Addresses of the American Philosophical Association* 79, no. 2 : 47–65.
- Fischer, John M., and Mark Ravizza. 1998. *Responsibility and control: A theory of moral responsibility*. Cambridge: Cambridge University Press.
- Fodor, Jerry A. 1983. *The Modularity of mind: An essay on faculty psychology*. Cambridge, Mass.: MIT Press.
- Frankfurt, Harry G. 1978. “The problem of action.” *American Philosophical Quarterly* : 157–162.
- Gallagher, Shaun, ed. 2011. *The Oxford handbook of the self*. Oxford: Oxford University Press.
- . 2001. “The practice of mind. Theory, simulation or primary interaction?” *Journal of Consciousness Studies* 8, 5-6 : 83–108.
- . 2005. *How the body shapes the mind*. Oxford: Clarendon Press.
- . 2009. “Two problems of intersubjectivity.” *Journal of Consciousness Studies* 16, 6-1 : 289–308.
- Gergen, Kenneth J. 2009. *Relational being: Beyond self and community*. Oxford: Oxford University Press.
- Gibson, James J. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gilbert, Margaret. 1990. “Walking together: A paradigmatic social phenomenon.” *Midwest Studies in Philosophy* 15, no. 1 : 1–14.
- Haidt, Jonathan. 2001. “The emotional dog and its rational tail: a social intuitionist approach to moral judgment.” *Psychological Review* 108, no. 4 : 814.
- . 2006. *The happiness hypothesis: Finding Modern Truth in Ancient Wisdom*. London: Arrow.
- Harman, Gilbert. 1999. “Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error.” *Proceedings of the Aristotelian Society* 99: 315–331.
- Hassin, Ran R., James S. Uleman, and John A. Bargh, eds. 2007. *The new unconscious*. Oxford: Oxford University Press.
- Heidegger, Martin. 1927/1968. *Being and time*. Translated by John Macquarrie and Edward Robinson. New York: Harper Perennial.
- Hobbes, Thomas. 1651/2016. *Leviathan*. London: Penguin Books.
- Huizinga, Johan. 1938/2016. *Homo ludens: A study of the play-element in culture*. Seattle: Angelico Press.
- Husserl, Edmund. 1989. *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy. Book 2. Studies in the phenomenology of constitution*. Translated by Richard Rojcewicz and André Schuwer. Collected works volume 3. Dordrecht: Kluwer Academic.
- Hutto, Daniel D., and Erik Myin. 2013. *Radicalizing enactivism: Basic minds without content*. Cambridge, Mass: MIT Press.
- Isen, Alice M., and Paula F. Levin. 1972. “Effect of feeling good on helping: cookies and kindness.” *Journal of personality and social psychology* 21, no. 3 : 384–388.
- Jaegher, Hanne de, and Ezequiel Di Paolo. 2007. “Participatory sense-making.” *Phenomenology and the Cognitive Sciences* 6, no. 4: 485–507.
- Joas, Hans. 1985. *Praktische Intersubjektivität: Die Entwicklung des Werkes von George Herbert Mead*. Frankfurt am Main: Suhrkamp.
- Johnson, Mark. 2014. *Morality for humans: Ethical understanding from the perspective of cognitive science*. Chicago: University of Chicago Press.
- Kane, Robert. 2007. “Libertarianism.” In *Four views on free will*. Edited by John M. Fischer et al., 5–43. Oxford: Blackwell.
- Kant, Immanuel. 1787/2009. *Critique of pure reason*. Translated by Paul Guyer. Cambridge: Cambridge University Press.
- Koffka, Kurt. 1935. *Principles of Gestalt psychology*. London: Kegan Paul.
- Korsgaard, Christine M. 2009. *Self-constitution: Agency, identity, and integrity*. Oxford: Oxford University Press.
- Lewis, J. D., and Richard L. Smith. 1980. *American sociology and pragmatism: Mead, Chicago sociology, and symbolic interaction*. Chicago: The University of Chicago Press.
- Margalit, Avishai. 1996. *The decent society*. Cambridge, Mass: Harvard University Press.
- McDowell, John H. 2007. “What myth?” *Inquiry* 50, no. 4: 338–351.
- . 2009. “Avoiding the Myth of the Given.” In *Having the world in view: Essays on Kant, Hegel, and Sellars*, 256–74. Cambridge, Mass: Harvard University Press.
- Mead, George H. 1934. *Mind, self and society from the standpoint of a social behaviorist*. Chicago: University of Chicago Press.
- . 1938/1972. *The philosophy of the act*. Chicago: The University of Chicago Press.
- Mele, Alfred R. 2009. *Effective intentions: The power of conscious will*. New York: Oxford Univ. Press.
- Merleau-Ponty, Maurice. 1942/1965. *The structure of behavior*. Translated by Alden L. Fisher. Pittsburgh: Duguesne University Press.
- Metzinger, Thomas. 2003. *Being no one: The self-model theory of subjectivity*. Cambridge, Mass: MIT Press.
- Milgram, Stanley. 1974. *Obedience to authority: An experimental view*. Enfield: Perennial.

- Nietzsche, Friedrich. 1886/2013. *On the genealogy of morals: A polemic*. London: Penguin Books.
- Nisbett, Richard E., and Timothy D. Wilson. 1977. "Telling more than we can know: Verbal reports on mental processes." *Psychological Review* 84, no. 3 : 231.
- Noë, Alva. 2004. *Action in perception*. Representation and mind. Cambridge, Mass: MIT Press.
- . 2005 "Against intellectualism." *Analysis* 65, no. 288 : 278–290.
- . 2012 *Varieties of presence*. Cambridge, Mass: Harvard University Press.
- O'Regan, J. Kevin. 2011. *Why red doesn't sound like a bell: Understanding the feel of consciousness*. Oxford: Oxford University Press.
- O'Regan, J. Kevin, and Alva Noë. 2001. "A sensorimotor account of vision and visual consciousness." *Behavioral and brain sciences* 24, no. 05 : 939–973.
- Pacherie, Elisabeth. 2011. "Self-Agency." In *The Oxford handbook of the self*. Edited by Shaun Gallagher, 442–64. Oxford: Oxford University Press.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Parfit, Derek. 2011. *On what matters*. Oxford: Oxford University Press.
- Pereboom, Derk. 2007. "Hard Incompatibilism." In *Four views on free will*. Edited by John M. Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas, 85–125. Oxford: Blackwell.
- Rietveld, Erik, and Julian Kiverstein. 2014. "A Rich Landscape of Affordances." *Ecological Psychology* 26, no. 4 : 325–352.
- Sartre, Jean-Paul. 1943/2012. *Being and nothingness: An essay on phenomenological ontology*. New York: Washington Square Press.
- Schechtman, Marya. 1996/2007. *The constitution of selves*. Ithaca: Cornell University Press.
- . 2014. *Staying alive: Personal identity, practical concerns, and the unity of a life*. Oxford: Oxford University Press.
- Schiller, Friedrich. 1795/2016. *On the aesthetic education of man*. London: Penguin Books.
- Searle, John R. 1990. "Collective intentions and actions." In *Intentions in communication*. Edited by Philip R. Cohen, Jerry L. Morgan and Martha E. Pollack, 401–15. Cambridge, Mass.: MIT Press.
- Sellars, Wilfrid. 1963. "Philosophy and the scientific image of man." *Science, perception and reality* 2 : 35–78.
- Sie, Maureen, and Arno Wouters. 2010. "The BCN challenge to compatibilist free will and personal responsibility." *Neuroethics* 3, no. 2: 121–133.
- Thompson, Evan. 2007. *Mind in life: Biology, phenomenology, and the sciences of mind*. Cambridge, Mass, London: Belknap.
- . 2014. *Waking, dreaming, being: Self and consciousness in neuroscience, meditation, and philosophy*. New York: Columbia University Press.
- Tomasello, Michael. 2016. *A natural history of human morality*. Cambridge, Mass.: Harvard University Press.
- Turvey, Michael T., and Robert E. Shaw. 1979. "The primacy of perceiving: An ecological reformulation of perception for understanding memory." *Perspectives on memory research*: 167–222.
- Uleman, James S. 2007. "Introduction." In *The new unconscious*. Edited by Ran R. Hassin, James S. Uleman and John A. Bargh, 3–15. Oxford: Oxford University Press.
- Varela, Francisco J. 1988. "Structural Coupling and the Origin of Meaning in a Simple Cellular Automation." In *The semiotics of cellular communication in the immune system*. Edited by Eli E. Sercarz, 151–61. Springer.
- Varela, Francisco J., Evan T. Thompson, and Eleanor Rosch. 1991. *The embodied mind: Cognitive science and human experience*. The MIT Press.
- Vosgerau, Gottfried, and Albert Newen. 2007. "Thoughts, motor actions, and the self." *Mind & Language* 22, no. 1 : 22–43.
- Vygotsky, Lev S. 1978/2007. *Mind in society: The development of higher psychological processes*. Harvard University Press. Cambridge, Mass.
- Wegner, Daniel M. 2002. *The illusion of conscious will*. Cambridge, Mass: MIT Press.
- Weichold, Martin. 2015. *Zwischen Reflex und Reflexion: Intelligenz und Rationalität im unreflektierten Handeln*. Berlin: de Gruyter.
- . 2016. "Mastering Sociocultural Contingencies: On Extending the Sensorimotor Theory to the Domain of Culture." *Journal of Consciousness Studies* 23, 5-6 : 203–227.
- . "Social Authenticity: Towards a Heideggerian Analysis of Social Change." In *From Conventionalism to Social Authenticity: Heidegger's Anyone and Contemporary Social Theory*. Studies in the Philosophy of Sociality 10. Edited by Hans B. Schmid and Gerhard Thonhauser. Forthcoming.
- Wilson, Timothy D. 2004. *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, Mass, London: Belknap.
- Wittgenstein, Ludwig. 1953/2009. *Philosophical Investigations*. 4th ed. Translated by G.E.M. Anscombe, P.M. Hacker and J. Schulte. Chichester: Wiley-Blackwell.
- Zahavi, Dan. 2005. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Cambridge MA: Bradford Book/MIT Press.
- . 2014. *Self and other: Exploring subjectivity, empathy, and shame*. Oxford: Oxford University Press, 2014.
- Zimbardo, Philip G., Craig Haney, W. C. Banks, and David Jaffe. 1973. "The mind is a formidable jailer: A Pirandellian prison." *The New York Times Magazine* 8: 38–60.