

Penultimate draft of Rowland Stout, “Being subject to the rule to do what the rules tell you to do”, from Weiss, B. and Wanderer, J. (eds.), *Reading Brandom*, Routledge, forthcoming.

## 1. Kant’s Problem

One way to start thinking about agency is to try to distinguish the special way that reasons are involved in action from the way that reasons are involved in inanimate nature. Consider the following pair of explanations:

Explanation A. *The reason the soufflé collapsed is that the oven door was opened at the wrong time.*

Explanation B. *The reason John collapsed onto the sofa was that he was exhausted after a hard day at work.*

The approach to agency that I am considering, one that goes back through Anscombe and Davidson to Kant and perhaps Aristotle, is to say that there is something special about the explanation in case B, and that it is this special ‘reason-giving’ quality of the explanation of John’s collapsing that accounts for it being a manifestation of his *agency* and an *intentional* action.<sup>1</sup> But, of course, this approach has to distinguish explanation B from explanation A in such a way that it is made clear why explanation A does not lead to talk of agency and intentional action too.

Each of these explanations may be expressed as an inference:

Inference A. *The oven door was opened at the wrong time; so the soufflé collapsed.*

---

<sup>1</sup> See Anscombe (1957) and Davidson (1980, essay 1).

Inference B. *John was exhausted after a hard day at work; so he collapsed onto the sofa.*

And in each case the entitlement to make the inference may be made explicit by a rule – albeit not a strict or universal rule:

Rule A. *If the oven door is opened at the wrong time when a soufflé is being cooked and various other conditions are met, then, other things being equal, the soufflé should collapse.*

Rule B. *If one is exhausted after a hard day at work and various other conditions are met, then, other things being equal, one should collapse onto the sofa.*

So the approach to agency that I am considering, which asks what is special about the way reasons are involved in action, might equally ask what is special about the way inferences or rules are involved in action. For example, the task as described by Kant (e.g. in the *Groundwork of the Metaphysics of Morals*, 36) is to distinguish the way that a rational agent's behaviour accords with rules from the way other things in nature accord with rules.<sup>2</sup>

It may seem that the difference between John's behaviour according with a rule and a soufflé's activity according with a rule is a difference in the nature of the rules themselves. The rule that a person acts in accordance with is a normative one that says what the person *should* do in certain circumstances, while the rule that the behaviour of the soufflé conforms with is not a normative one and just says what the

---

<sup>2</sup> The page numbers for the *Groundwork* are from the second edition and for the *Critique of Practical Reason* are from the Royal Prussian Academy of Sciences 1929 edition edited by Natorp.

soufflé *will* do in certain circumstances. A law of nature is one kind of conditional claim and a normative rule is another.

But while it is clear that the rules that guide reasonable action are normative and the laws of nature are not, it is not clear whether this difference is *intrinsic* to the rules or depends on the sort of relationship the rules have with the things they apply to. Indeed it is difficult to see how to explain the idea that the normativity or factuality of a rule is an intrinsic feature of the rule rather than having something to do with what is done with the rule – with how it is used. For example it does not seem sufficient just to say that normative rules use normative vocabulary and factual ones do not. One can state normative rules without using words like “ought” and “should” and equally one can use these words in stating factual rules (as with my Rule A for soufflés above).<sup>3</sup>

So I will work on the assumption here that the difference between a normative and a factual rule is a difference in the sort of relationship each has with the things it applies to. The idea roughly is that a law of nature *describes* a natural process, whereas a normative rule *guides* a process of acting. In according with a normative rule a rational agent like John is *sensitive* to the rule. In according with a law of nature a soufflé is merely *subject* to it.

In setting up the issue in this way I am following the line of Robert Brandom (1994, 30 ff.) in his discussion in chapter 1, section IV of *Making It Explicit*. Brandom endorses Kant’s way of marking the distinction between the way rational agents and inanimate objects accord with rules, which is the following: “Everything in nature works according to laws. Only a rational being has the power to act according to the idea of laws – i.e. according to principles.” (*Groundwork*, 36.)

---

<sup>3</sup> It is often claimed that the normative rules that guide action do not say what *should* be done but only what *may* be done. This looks like an intrinsic difference between normative rules and factual rules after all. But whatever we say about the possible role for rules of entitlement in reason-giving explanation, all that I require here is that *sometimes* there is a role for rules of obligation

Brandom reads this as meaning that rational agents act according to their own conception of rules. And this in turn is taken to mean that rational agents act according to some attitude they have towards rules – namely the attitude of acknowledging these rules.

[T]he point he is making is that we act according to our *grasp* or *understanding* of rules. The rules do not immediately compel us, as natural ones do. Their compulsion is rather mediated by our *attitude* towards those rules. What makes us act as we do is not the rule or norm itself but our *acknowledgment* of it. It is the possibility of this intervening attitude that is missing in the relations between merely natural objects and the rules that govern them. (Brandom 1994, 31)

My task here is the same as one of Brandom's – to make proper sense of the difference between a rational agent's sensitivity to rules and an inanimate object's subjection to rules (or to reasons). I will not present any more argument for the shared presumption that sensitivity to rules is essential for intentional agency, nor for Brandom's further claim that this in turn is essential for intentionality of any sort.

In the next section I argue that Brandom's introduction on Kant's behalf of an intervening attitude between the rules and the agent's behaviour in order to explain what is involved in sensitivity to rules is unhelpful and suggests an unsatisfactory picture of action. Then I present an alternative way to understand sensitivity to rules, one which makes no use of the problematic idea of intervening attitudes but which at the same time does justice to Kant's claim that a rational agent acts according to the idea of rules. I will propose that, while the working of processes generally is described by *using* rules, the working of a process of acting is described by *referring*

to rules – or more fully it is described by using one rule which refers to others. I take this to be a very natural way to understand the idea of sensitivity to rules. What distinguishes normative rules from factual rules is simply that, rather than describing any actual process, they are referred to by the rules that do describe actual processes. So, processes that involve agency are described by second-order rules – rules that refer to other rules. Where the behaviour of all natural things is subject to rules, the behaviour of agents is subject specifically to the rule to do what the rules tell you to do.

In this way agents not only act according to rules, they act according to the idea of rules. I think that what I have presented here is pretty close to Kant's own conception of rational agency. If you act from duty you act out of respect for the law, according to Kant, and this is doing what you do because this is what the rules tell you to do.

## 2. Brandom's approach

Kant's conception of the will as a sort of causality that belongs to living things in so far as they are rational is spelt out in the *Groundwork of the Metaphysics of Morals* and in the *Critique of Practical Reason* as the power to act according to the idea of rules or laws. It is described as “das Vermögen, nach der Vorstellung der Gesetze, d.i. nach Prinzipien, zu handeln” in the *Groundwork*, 36 and very similarly in *Groundwork*, 63. And in the *Critique*, 32 he uses a similar formulation but talks about rules instead of laws, when he describes the will as “ein Vermögen ... ihre Kausalität durch die Vorstellung von Regeln zu bestimmen.”

In this paper I will, like Brandom, make nothing of the distinction between rules and laws and indeed talk of them interchangeably. What I want to focus on instead is how to treat the word ‘idea’ or ‘Vorstellung’ here. Should we understand the *idea* of laws or rules to be the agent’s own representation of them - an aspect of their mental life? Or should we understand the *idea* of laws or rules to be an abstract impersonal representation of them - like “the representation of chivalry in fourteenth century poetry”? The issue is rather prejudiced by Paton’s 1948 translation of the *Groundwork* where “*der Vorstellung der Gesetze*” is translated as “*his idea of laws*” (my italics).

Brandom treats the idea of rules as the agent’s own attitude towards the rules – their attitude of acknowledging the rules. So Brandom’s reply to Kant’s problem is to say that in the case of rational agents, norms or rules are first brought into the agent’s motivational system in virtue of the agent acknowledging them. Rational action is the causal response to these acknowledgements. Brandom is not here claiming that rules and reasons are themselves attitudes or are about attitudes. His claim is that there must be attitudes causally intervening between the rules and the actions.

There is often a temptation in philosophy when faced with the challenge of explaining what is special about the relation between A and B to introduce a third element C and say that what is special is the existence of this intervening element. This is usually an unhelpful move since it takes the initial problem of characterizing one relation and exchanges it for the two problems of characterizing the relation between A and C and characterizing that between C and B.

A classic example of this is the attempt to explain what it is for a person to be consciously aware of something. No progress has been made by saying that the person is in a certain relation of awareness with some internal object and that the

internal object is in the relation of representation with the external thing. Indeed it looks to many as if the problem is now quite insoluble once this first move is made.

Brandom acknowledges that Kant's strategy as he interprets it does have something in common with Descartes' invocation of intervening representations in explaining the possibility of error about external things (Brandom, 1994, 31-2). But Brandom's own version of this strategy is distinctive, since he takes the relevant intervening attitude to belong to a specifically *linguistic* social practice, thus according explanatory priority to linguistic practice over rational agency. The social nature of the attitudes intervening between rules and behaviour means that Brandom is not endorsing Descartes's extraordinary strategy of going inwards to explain our special relationship with the outer world.

Nevertheless I think that Brandom's attempt to explain the process of acting according to reason as a two-stage process of acknowledging a rule and then responding to that acknowledgment can give us a distorted conception of the causal process that constitutes action. The distortion arises if in the two-stage model the rationality characteristic of agency is confined to the first stage – the production of attitudes. It looks like this must be the case because it is only in this first stage that the rules themselves are involved. In the second stage it is the agent's attitudes to the rules not the rules themselves that are involved in the process. Since the agent's attitudes to rules (or reasons) are not themselves rules (or reasons) it looks as if the transition from these attitudes to things actually being made to happen in the world around does not itself involve the rationality characteristic of agency.

In this problematic version of the two-stage model of action, the first stage involves sensitivity to rules (or reasons), but does not involve anything actually happening, and the second stage involves things actually being made to happen but involves no sensitivity to rules (or reasons). This fails to take seriously the idea of

action as a process of rationally transforming the world – i.e. a process in which the changes characteristic of the action involve the rationality characteristic of agency. Instead the rationality characteristic of agency is manifested in the production of attitudes; the transformation of the world characteristic of the action is not taken to be a manifestation of rationality in action, but rather a response to such a manifestation.<sup>4</sup> Intervening attitudes between the rules and the action which are supposed to explain the special relationship between rules and action serve merely to separate the rules from the action.

Now Brandom might deny that the rationality characteristic of agency is confined to the first stage of his two-stage model and accordingly deny that the second stage in his two-stage model is just a causal response to the acknowledgement of rules. He might say that the first stage involves forming attitudes and the second stage involves *acting* according to those attitudes. This seems quite reasonable, albeit not a move he actually makes. He might then deny that acting according to attitudes should be understood simply as a causal process that takes those attitudes as inputs. Instead the process of acting according to attitudes – the second stage – itself involves something characteristic of agency; this would be something over and above the special nature of the inputs to that process.

The trouble with this from Brandom's point of view is that he would lose the supposed explanatory gain from introducing intervening attitudes between rules and the world in his attempt to answer Kant's problem. What distinguishes a rational agent's acting being guided by rules from an inanimate object's behaviour being subject to rules would not then be just the existence of an intervening attitude in the

---

<sup>4</sup> Some philosophers have accepted that the work of agency is over before the body even starts to move. For example, Pietroski (2000), developing Hornsby's (1980) identification of acting with trying, argues



rational agent's process of acting. There is something about the process of acting in addition to the mere existence of an intervening attitude that explains why this process manifests rational agency, and Brandom has not told us what it is. In this case Brandom has not answered Kant's problem.

### 3. Subjection to second-order rules

The key stage in Brandom's derivation of the two-stage model of action is his interpretation of Kant's phrase: "acting according to the idea of rules". Brandom takes it to mean: "acting according to the agent's own idea of the rules." But there is another way to read it.

Consider again the distinction between being subject to something and being sensitive to something. In being disposed to rust in the presence of moisture in the air the behaviour of a bit of iron is *sensitive* to the presence of moisture. At the same time it is *subject* to the rule that if there is moisture in the air, rust should happen. But although the bit of iron responds to the presence of moisture in the air it does not respond to the presence of this rule. Its behaviour is not *sensitive* to the rule that if there is moisture in the air then rust should happen.

We might think of the rules of physical science in two ways – either as the rules of physical science as it now stands or as the rules of an ideally completed physics. In the former case it is clear that the rules that figure in physical science as it stands are determined by the behaviour of physical processes and not the other way around. If physical science included a different specification of the rust rule – say that

---

that actions are *inner causes* of bodily movement. But I take it that this is a position only to be adopted if one has painted oneself into a corner, and my task here is to see if we can step out of that corner.

iron should not rust in the presence of moisture but should rust otherwise – it would have no effect on the piece of iron’s behaviour. On the other hand, if the rules of completed physics (assuming that this idea makes any sense) had this different rust rule then certainly in that possible world the behaviour of the iron would be different to the way it actually is. But again it would not be adapting to the presence of this different rule.<sup>5</sup>

A rational agent on the other hand – one who is guided by rules - has the capacity not only to be sensitive in their behaviour to the features of the world that rules describe, but also sensitive to the rules themselves. A rational agent can respond not only to the presence of moisture in the environment, but can also respond to what the rules tell it should be done in such a situation. If the rules change, the behaviour of a rational agent changes. This is something that Brandom’s two-stage model was supposed to capture.

For this idea of sensitivity to rules to make sense we must think of the rules as existing somehow. The sensitivity at issue is not in the first instance sensitivity to the *truth* of rules, but is sensitivity to the *presence* of rules. Rules may be present explicitly in some rule-book or in some scientific theory. But they may also be present implicitly in some normatively structured set of conventions. We may also think of rules as being present even more implicitly in that set of rules that *should* structure a set of conventions.

On the view I am recommending here, the behaviour of rational agents is still subject to a rule – a second-order one that makes reference to other rules. The behaviour of a rational agent is subject to the rule, “Whatever those first-order rules say should happen should happen.” But subjection to this rule is not a separate

---

<sup>5</sup> This point is developed in section 4 below.

process coming after the real work of rational agency. Subjection to this second-order rule just is sensitivity to reason.<sup>6</sup> For the piece of iron, subjection to the rust rule just is sensitivity to the presence of moisture. In the same way, for a rational agent, subjection to a second-order rule just is sensitivity to the presence of first-order rules. This makes some sense of the distinction between acting according to rules and acting according to the idea of rules. As with the two-stage model, acting according to the idea of rules is having one's behaviour causally determined at one stage removed from the normative rules. But this is not to be understood as having one's behaviour causally determined by something else that is the true rational response to the normative rules - for example the attitude of acknowledging the normative rules. Instead it is to be understood as being causally determined by a second-order rule that demands sensitivity to what the first-order rules say. The first-order rules determine behaviour through the idea of them in the second-order rules. The *idea of rules* is taken here not to be the presentation of the rules in the mind of the agent but rather the presentation of the rules that figures in the second order rule that the agent's behaviour is subject to.

I think that this notion of subjection to properly second-order rules is the right way to make sense of Kant's conception of reverence or respect for the moral law (introduced in the *Groundwork* and developed in chapter 3 of Book 1 of the *Critique of Practical Reason*). Kant regards this attitude of acknowledging the subordination of one's will to the law as essential to truly rational activity. But he explicitly denies

---

<sup>6</sup> This is reminiscent of Sellars's (1963, 325) suggestion when worrying about Kant's problem himself in "Some Reflections on Language Games" that "learning a game involves learning to do what one does *because doing these things is making moves in the game.*"

that reverence for the moral law mediates our moral agency. It is not a component in the process of acting rationally.

Immediate determination of the will by the law and consciousness of this determination is called ‘reverence’, so that reverence is regarded as the effect of the law on the subject and not as the cause of the law. (*Groundwork*, 16)

Doing what the law says because it is the law counts as reverence for the law or respect for the law. Doing what the law says because it is the law is simply being subject in one’s behaviour to the second-order rule: what the law says should be done should be done. The attitude of respect for the law does not have to be conceived as an independent input in law-guided activity. Instead, to have that attitude *just is* to behave in a way that is sensitive to the law and subject to the second-order rule.

The attitude of respect for the law does not have explanatory priority over rational agency. Once we have a conception of rational agency as subjection to second-order rules, we can attribute the attitude of respect as a necessary consequence of rational agency. For Kant, respect for the moral law is a feeling that can be attributed to moral agents *a priori* (*Critique of Practical Reason*, 73) – simply in virtue of their being moral agents. Equally, the attitude of acknowledging the rules required by Brandom may be taken to be necessary to rational agency, but as a consequence rather than a mediating component.

So, you would not have the rational sensitivity I am describing unless you had the attitude of respect for these rules. But my suggestion is the behaviourist one that such an attitude is not a *mediating component* in rule-guided behaviour, but can be attributed to the rational agent *in virtue of* such rational sensitivity.<sup>7</sup> The attitude of

---

<sup>7</sup> For a development of the behaviourist suggestion, see my 2006.

respect for a rule or acknowledgement of a reason turns out to be necessary when a rational agent acts according to that rule or reason, but it does not figure in an account of what it is for a rational agent to act according to that rule or reason.

#### 4. Causal Rules

The viability of the distinction between subjection to rules and sensitivity to rules depends on one's view of causation. For example, if one held a Humean regularity theory, then one would say that rules of nature simply describe what regularly happens. Given this, the notion of *subjection* would be quite out of place – a kind of anthropomorphism - and Kant's problem would not get a grip. On this view there might still be room for a distinction between first-order and second-order rules of nature, but it could not be motivated in the same way.

I am assuming in the way that I have set up Kant's problem that there must be more to causation than this, and that things really do make other things happen. I am working with a picture of causation according to which causal processes result in other things happening, and rules of nature describe how these causal processes work. When something happens that results from some process, it is subject to the rules that describe how that process works. If a rule says that in such-and-such a circumstance such-and-such should happen, and this describes a causal process that does result in such things happening, then what happens is subject to that rule. At the same time it is sensitive to the circumstances that figure in the rule.

Applying this to the collapsing soufflé we can say the following sensible things. The soufflé's behaviour is sensitive to whether and when the oven door is

opened, and it is subject to the rule: if the oven door is opened at the wrong time and various other conditions are met, then, other things being equal, the soufflé should collapse. This rule describes the causal process that results in the soufflé collapsing. In the same way, the rust rule describes the causal process of rusting that results in the behaviour of the piece of iron. The inputs to this process, for example the presence or absence of moisture in the air, are what the behaviour resulting from the process is sensitive to.

Combining this picture of causal processes with the suggested response to Kant's problem gives the following view of rational agency. Being an agent - transforming the world according to reason - just is having one's behaviour result from a certain sort of process. The way that process works is described by a second-order rule - one which itself makes reference to rules. The process resulting in such behaviour that is sensitive to rules is the process of acting rationally.

It might be objected against this simple strategy for solving Kant's problem that the distinction between second-order rules and first-order rules must collapse. Whenever something is subject to a rule is it not automatically subject to a second-order rule recommending that it do what the first rule says? Consider again the piece of iron whose behaviour is subject to the rule that says that if there is moisture in its environment, rust should happen. It is equally subject to the second-order rule that says whatever that first-order rule says should happen should happen. If the recommendations of the first-order rule are rigidly fixed, then the two rules are equivalent.

It is clear that reference to a rule within a rule is quite redundant in this case. So we should identify sensitivity to rules with subjection to *properly* second-order rules. A rule is *properly* second-order only if it describes some sensitivity that cannot be described by any first-order rule. You get this sensitivity if the recommendations

of the rule that is specified in the second-order rule are not completely fixed by that specification.

Consider a law-abider. This person's behaviour is subject to a second-order rule that says that whatever the law says should be done should be done. If we could construct a first-order rule that was in fact the complete and final specification of the law, then subjection to this first-order rule would give exactly the same sensitivity as subjection to the second-order rule. But there can be no complete specification of the law. For every finite specification of the law it is always possible to think of circumstances not included in that specification. The recommendations for those circumstances are not determined in advance. So there are always alternative possibilities available to the law - it is open-ended. Also the law changes, and as it changes so does the behaviour of the law-abider. So whatever first-order rule we come up with will not quite capture the flexibility of the second-order rule that requires no complete first-order specification.

It might be responded that the laws of nature are open-ended in a similar sort of way. Perhaps the process of scientific discovery is by its nature unbounded and that any laws derived will only be *pro tem*. If the laws of nature are *pro tem* in the sense of not all being true, then natural events are not subject to the second-order rule that says that what should be done is what the laws of nature say should be done, and this objection fails. So let us assume that the laws of nature are *pro tem* in the sense of being unspecifiable in full at any one time.

But even if this is the case, it does not mean that the rule that says that whatever the laws of nature say should be done should be done is properly second-order. I take it that the rules of nature are simply the descriptions of the natural processes that result in things happening. These rules would be unspecifiable in full because no complete description of every process is possible. But in that case, to

claim that the rules of nature say that something should happen is precisely to claim that there is a natural process that results in it. So the second-order rule that says that whatever the rules of nature say should happen should happen is equivalent to the rule that says that whatever results from natural processes should happen. This rule is both empty and first-order.

To reiterate, my suggestion is that rational rule-guided behaviour is distinguished by the fact that it results from a process described by a properly second-order rule. This requirement that the rules that a rational agent's behaviour is subject to be properly second-order corresponds to a rejection of what Brandom calls 'regularism'. Regularism is the claim that the normative rules that an agent's behaviour is sensitive to can be read off from regularities in the way that agent (or others) behaves – "a norm implicit in practice is just a pattern exhibited by behaviour" (1994, 28). In my terms this is the claim that the normative rules that an agent's behaviour is sensitive to can be read off from the rules that their behaviour is subject to. If this were the case then the rules that their behaviour was subject to would not be properly second-order.

But my suggestion also counts as an example of regularism in a different weaker sense which seems also to be employed by Brandom. It does so in as much as it treats the rule-guidedness of behaviour as a feature of what is done – that what is done conforms with some specific regularity (1994, 63). The specific regularity in what is done that I am identifying is conformity to a second-order rule. Brandom argues that instead of explaining what it is for behaviour to be rule-guided in terms of what is done, one should explain it in terms of attitudes towards what is done – attitudes that are themselves implicit in practice.

But the regularity that my account appeals to – namely the regularity of conforming with a second-order rule – is not a regularity that can just be read off from



the behaviour of agents. This is because the first-order rule that is referred to is not explicitly *specified* in the second-order rule. So the distinction between treating a performance as subject to normative assessment and treating it as subject to physical laws is not obliterated on this account.

This is also why Brandom's central argument against regularism – Wittgenstein's gerrymandering argument – does not work against my account. According to this central argument, different regularities can always be found to match any actual history of performances, so that for any hypothetical further performance it conforms to one regularity and not to another. Wittgenstein illustrates this with the example of the 'Plus 2' rule. This means that the existence of a regularity does not have any normative significance since it provides no way to accept or reject future performances.

My account makes no attempt to explain normative significance in terms of the mere existence of a regularity. What determines correctness or incorrectness of a future performance is the first-order rule referred to in the second-order rule that does describe the behaviour. The first-order rule exists independently of the pattern of behaviour.

## 5. Brandom's Pragmatism

Brandom's aim in the philosophy of language is to account for representational concepts like truth, reference and existence in terms of structural features of the proprieties of linguistic inferences. True to his aim of constructing an 'inferentialist' as opposed to a 'representationalist' account of intentionality, Brandom treats the idea of *the idea of rules* in inferential rather than representational terms. He starts with the

idea that I have been questioning, namely that a rational rule-guided agent must have a certain attitude of acknowledging the rules mediating the rules and their behaviour. Then he explains this sort of attitude in terms of having a disposition to sanction, positively and negatively, correct and incorrect applications of the rule (1994, 35). This sanctioning behaviour can only be understood as a social practice. So, rational rule-guided behaviour (and hence intentionality itself) can only be understood in terms of social practices. Brandom presents the example of beating someone with sticks when they go wrong as a primitive example of this sort of social practice. But in the end his account of linguistic scorekeeping practices is going to be required to explain fully conceptual intentional states and behaviour (Brandom 1994, chapter 3).

My alternative suggestion is that what it is for an agent to act according to the idea of rules is for those rules to be referred to in the second-order rule that the agent's behaviour is subject to. Introducing such talk of reference at this stage of the story spoils the purity of the inferentialism, and Brandom certainly seems to be trying to avoid doing so. But it does not do this by saying that we have to understand the notion of representations in the mind to understand intentionality. It does this instead by saying that we have to understand the notion of one *rule* representing another in order to understand intentionality. Only a very reductive sort of inferentialism would be threatened by this.

So I am not sure to what extent my suggestion counts as a fundamental difference with Brandom. Brandom's goal in the section of *Making It Explicit* that I am looking at is to show that rational rule-guided behaviour (and hence intentionality itself) only makes sense in the context of social institutions of treating performances as correct or incorrect. My argument that rational rule-guided behaviour requires behaviour to be subject to second-order rules in no way obviates the need for some sort of account of how the rules that are referred to in these second-order rules are

supposed to exist. Brandom's Wittgensteinian argument against what he calls 'regulism' – not to be confused with 'regularism' – is supposed to show that no account of the existence of rules that relies entirely on their existing in a fully explicit way can work. Given this it is natural to try to make sense of the idea of sensitivity to rules by thinking of these rules (the ones that are embedded in second-order rules) as existing implicitly in social practice.

So I think there are two arguments we can extract from this chapter for the conclusion that rule-guided behaviour depends on the existence of a social practice, one starting with Brandom's interpretation of Kant's "acting according to the idea of the rules" and one starting from mine. Although Brandom does not endorse the starting point of this second argument, he endorses the rest of it. Let me call them the Official Argument and the Unofficial Argument.

Here is the Official Argument:

1. Rational rule-guided behaviour depends on the agent having normative attitudes to performances; this is Brandom's interpretation of "acting according to the idea of the rules".
2. The agent's normative attitudes to performances are dispositions to sanction, positively or negatively, correct or incorrect performances.
3. Such sanctioning behaviour only makes sense in the context of social practice.
4. So rational rule-guided behaviour (and hence intentionality) only makes sense in the context of social practice.

Here is the Unofficial Argument that uses my first premise instead:

1. Rational rule-guided behaviour depends on the agent being subject to a rule that refers to rules; this is my interpretation of “acting according to the idea of rules”.
2. A rule cannot be referred to by another rule unless it exists; it follows from this and 1. above that rule-guided behaviour depends on the *existence* of rules to be guided by.
3. Such rules cannot exist only in a fully explicit way but must be taken partly to exist implicitly in practice; this corresponds to Brandom’s rejection of regulism.
4. But such rules must exist independently of the agent’s own behaviour; this corresponds to Brandom’s rejection of regularism and follows in my account from the requirement that the rules that the agent’s behaviour is subject to are properly second-order.
5. A rule which partly exists implicitly in behaviour but is at the same time independent of an agent’s actual behaviour must depend on some other behaviour; in particular, it depends on the existence of a social practice of treating performances as correct or incorrect.
6. So rational rule-guided behaviour (and hence intentionality) only makes sense in the context of social practice.

I am not setting out here to defend the Unofficial Argument, just to present it as perfectly in line with what Brandom himself says apart from the first premise. The premise in stage 3 coincides with Brandom’s rejection of regulism. Regulism is taken to be the claim that norms can be fully understood in terms of explicit rules. Brandom approves of Wittgenstein’s regress of rules argument against this in the *Philosophical Investigations*.

Rules do not apply themselves; they determine correctnesses of performance only in the context of practices of distinguishing correct from incorrect applications of the rules. To conceive these practical proprieties of application as themselves rule-governed is to embark on a regress. Sooner or later the theorist will have to acknowledge the existence of practical distinctions between what is appropriate and what is not, admitting appropriatenesses according to practice as well as according to rules or explicit principles. (Brandom, 1994, 20-1)

The rejection of regulism combines powerfully with the rejection of regularism. According to the rejection of regulism, the rule whose existence an agent's behaviour is sensitive to cannot exist in a completely explicit way, but must be partly implicit in behaviour. And according to the rejection of regularism, the behaviour that the rule exists implicitly in is not the agent's own behaviour. This gives the conclusion that for some behaviour to count as rule-guided, there must be some *other* behaviour. Stage 5 in this argument is just the assumption that this other behaviour can only be a social practice. There is really nothing else for it to be, and since Brandom is keen to argue that norms are instituted by attitudes within a social practice of treating performances as correct or incorrect, he should have no problem with this. So if he accepted my first premise I think he would have a reasonable argument for the conclusion he is aiming at, one which despite its first premise contains much that he in fact endorses.

I want to end by illustrating this conception of rational agency as subjection to second-order rules using what must be one of the most basic skills in the progression to fully-blown rational agency - namely imitation. I present this as an alternative to Brandom's story of primitive sanctioning behaviour – beating wrongdoers with sticks. Brandom takes it to be a condition of adequacy of his enterprise that it make “the advent of the favored sort of linguistic intentionality less mysterious. ... The hope is that doing so will offer guidance concerning what would be involved in diagnosing aliens as exhibiting such states, and programming computers or teaching merely sentient animals to exhibit them” (1994, 7). Brandom's suggestion seems to be that teaching merely sentient animals to punish and reward each other would get them going on the right path. I suggest that teaching them to *imitate* sapient creatures is what is required instead.

Roughly speaking, the imitator's behaviour is subject to the rule: “Whatever she does you should do too.” Whether or not this is a second-order rule depends on how we take the notion of doing something. At its most basic the rule is to be understood as something like the following: “Within some repertoire of possible movements, whatever movements of her body are happening should happen in your body.” To be subject to this rule is to be sensitive to her movements and to mirror them. It may be that the behaviour of dolphins in play or human babies in mirroring an adult's facial expression is subject to this rule. By itself, it involves no sensitivity to rules.

But, if the imitator goes beyond repetition of mere bodily movement patterns to repetition of *ways of behaving*, it becomes genuinely subject to a second-order rule. Then the rule is to be understood as: “If her behaviour is subject to a certain rule then you should do what that rule recommends.” Babies develop from mirroring

movements to mirroring ways of behaving. If I am looking at or grasping an object, the baby may look at or grasp the same object. This involves producing quite different movements to mine, but movements which are subject to the same rules that my behaviour is subject to. This gets clearer still when the young child is engaged in play-acting.

But the beginning of real social practice comes when the rule you are imitating is itself a second-order rule. At this point you are not just simulating someone's sensitivity to certain conditions, you are simulating her sensitivity to rules. You have become a social conformer. A social conformer starts off being subject to the rule that can be expressed as follows: "Follow the rules that she is following." But since the *she* in question is also following the rules that someone else is following, and so on, they are all also subject to the rule: "Follow the rules that *they* are following."

"They" in this rule ranges over some group of people. There are indefinitely many special cases of this rule depending on the social group that figures in it. For example, proper members of this family or community are subject to the rule: "Follow the rules that members of this family/community follow." Good soldiers are subject to the rule: "Follow the rules that soldiers follow." Polite people are subject to the rule: "Follow the rules of polite society," and so on.

Real agents are not usually *merely* subject to such rules. Being subject to such rules usually involves sensitivity at a higher level. You follow such rules for a *reason*, and in doing so you are subject to a higher-order rule that demands sensitivity to that reason. But at an early stage of the progression towards full rationality, I think it is possible to be subject to such rules for no reason. You have been brought up in a certain way, and this is how you behave. Your initiation in the practice may not have been a matter of rational choice. What Gibbard (1990, chapter 4) calls "being in the

grip of norms” should be understood in the philosophy of action in advance of understanding the conscious acceptance of norms.

This makes clear how far away from Kantian autonomy we are at this stage of the progression. Something a bit like autonomy is present in all second-order rule-guided behaviour. Being subject to properly second-order rules involves not just responding to the recommendation of rules, but also working out what these recommendations are. But working out what the recommendations are is not the same as determining what they should be. This is clear from considering the law abider. They have to work out what the law is; but of course they do not determine it in the sense required for autonomy.

Subjection to the rule: “Follow the rules that they are following,” can move to something a bit closer to autonomy if membership of the group that “they” ranges over is determined by the fact that this is the very rule that they are following. We then get subjection to the rule: “Follow the rules that this sort of rule-follower follows.” In subjection to that rule you automatically become one of them, and are subject to the satisfyingly holistic second-order rule: “Follow the rules that *we* are following.” Being a legislating member of the kingdom of ends does not seem so far away after all.

Whether fully autonomous rational agency can be made intelligible as part of this progression from mere imitation remains to be seen. But these examples in the second half of the progression from imitation are certainly cases of acting for reasons. That they are such can be made intelligible without appeal to any motivating role for the attitudes the agent has to the rules they are following.



## REFERENCES

- Anscombe, G. E. M. 1957. *Intention*. Oxford: Blackwell.
- Brandom, R. 1994. *Making It Explicit*. Cambridge Mass.: Harvard University Press.
- Davidson, D. 1980. *Essays on Actions and Events*. Oxford: Oxford University Press.
- Hornsby, J. 1980. *Actions*. London: Routledge.
- Kant, I. 1948 edition. *Groundwork of the Metaphysics of Morals*. tr. Paton, H. as *The Moral Law*. London: Hutchinson.
- Pietroski, P. 2000. *Acting and Causing*. Oxford: Oxford University Press.
- Sellars, W. 1963. *Science, Perception and Reality*. London: Routledge.
- Stout, R. 2006. *The Inner Life of a Rational Agent*. Edinburgh: Edinburgh University Press.