Published
Louis Caruana, 'Realism and Rule-following', in: *Normatività Fatti, Valori*, R. Egidi, M. dell'Utri, & M. De Caro (eds) (Macerata: 2003), pp. 143-152.

Realism and Rule-Following

Louis Caruana SJ
Gregorian University

In the most general sense, realism can be described as the philosophical position according to which there is a mind-independent world, and moreover we humans can come to know something about it. In such a general form, this position has found applications in various areas, such as the question of universals, the question of scientific discourse, the question about moral predicates, the explanation of perception, and the question of evidence-transcending truths. In this paper, my general aim is to offer at least one satisfactory way of seeing how realism can also be crucial in understanding rule-following.

I will be understanding rule-following as an activity that involves the intention of being consistent in keeping to some standard or norm. Examples of such activity occur in the use of words, the movement of chess pieces, the adoption of certain behaviour patters that are considered acceptable and thus good, and the avoidance of others considered bad. Most of the work done in explaining rule-following, in the tradition of analytic philosophy, has been in the area dealing with the use of words, the area called semantic normativity. My strategy will be to start from the main results obtained up to now in this area of semantic normativity and then try to apply them to other areas, the most important of which is the area of the behaviour of people who, in general, desire to do good and avoid evil.

First, one must recall some fundamental aspects of rule-following. Being rational involves being able to treat something as a rule. This, however, can only be done if some conditions are satisfied regarding the nature of rules. The most intriguing condition is that the rule in question must be a normative constraint over an

indefinite variety of cases, so that one can determine the correct response in each of those cases. The well-known problem here is to see how something that is normative over an *indefinite* number of cases can be suitably accessible to a finite mind.

In trying to solve this problem in the area of semantic normativity, the normal procedure is to consider a very simple arithmetical rule as a model of semantic rules. Take the rule of addition. It can be shown that, if we start with a finite number of typical cases of addition, with the idea of showing what addition means, then nothing can guarantee that novel cases should follow suit. In other words, any rule identifiable from a finite number of cases is underdetermined. Other rules may be consistent with the finite number of cases and yet diverge at some future point from the usual addition rule we are familiar with (Wittgenstein 1953, § 197-202; Kripke 1982; Boghossian 1989; Baker and Hacker 1985; Holtzman and C. Leich 1981; Pettit 1990; Wright 1984).

So the question is: once I go beyond the domain of the exemplary cases, what is it that keeps me on the right track? What makes me stick to a rule and not extend the application to new cases in an irregular way? The initial procedure of rule-following is quite clear. There are two steps. First, I take certain cases to exemplify a particular rule; second, I extrapolate from the rule towards new cases by letting the inclination that has been thereby generated lead me on. These are not problematic steps. What causes concern is what follows. There must be something that further distinguishes between following the rule rightly from following it wrongly. If I cannot in principle be ever mistaken in following a rule beyond its exemplary cases, then there can be no rule-following. This is certainly counter-intuitive because we do find ourselves following various rules with considerable success. So, an account is needed to explain why people are pushed to extrapolate consistently in one direction and not in another, even after they run out, as it were, of the exemplary cases.

Following the lead of Wittgenstein, some philosophers have argued that the criterion of correctness, at least as regards meaning, is derived from what members of the meaning-generating community actually do, in other words, the criterion is derived from the practice (e.g. Kripke 1982). One way of understanding their main argument is in terms of explanation. Consider the following explanation:

> This group is behaving in this regular way because it is following rules R_1, R_2, R_3, ...

Is this an acceptable explanation? It seems not. On closer scrutiny, one realises that the explanation is vacuous. We give the impression we are explaining the regular behaviour in terms of rules, forgetting that our concept of rule is itself understandable only in terms of regular behaviour. If asked what it means to act according to rule R_1, we have nothing else to say than that to act according to rule R_1 means doing something in a regular way. If we repeat this for R_2, R_3, and the rest, we see that the explanation is just circular. We are effectively saying that

> This group is behaving in this regular way because it is behaving in a regular way.

What we should do is have the explanation go the other way round. We should not explain the practice by alluding to rules but explain the occurrence of rules by referring to the practice. A practice within a group should not be seen as a *product* of rule-following. On the contrary, a social practice is best seen as a *generator* of a set of rules, rather than the result of following them. One may concede here that, to some extent, rules have a limited but significant role as 'causes' of practice. This occurs when 'new-comers' to a group need socialisation to be initiated in the practice under consideration. This role of rules however should not be exaggerated and given absolutely priority, as if it were indispensable also to those established members of the group who are in no need of initiation. The major conclusion remains therefore that the ultimate source of normativity in rule-following is located in the community as it engages in a given practice.

To illustrate this fundamental point, it is useful perhaps to mention how Wittgenstein extended this idea in order to give an original account of mathematics. For him, a mathematical proof is not a novel application of established rules, nor a discovery of hidden mathematical structures, but a creative way of changing a common inclination among the community of specialists (Wittgenstein 1979a IV §30, VII §9; 1979b p. 185-6). This change of inclination involves the act of convincing others of something they had never imagined before. The mathematician originator of a new mathematical proof or concept has the task of doing something that will convince his colleagues that what he does is essentially supplying a proof. For example, when mathematicians proved that denumerable infinity, $\aleph_0$, is a *number*,

they went through a procedure that convinced their colleagues that they have indeed supplied a proof.

The proof itself is part of the novelty. For Wittgenstein, mathematical proofs are taken as paradigms within which a particular language-game is possible. Not everything is allowed, however. In fact, for this particular example, he thinks mathematicians are mistaken whenever they use the expression 'the number of elements in an infinite series' as if it had a clear meaning. The use of $\aleph_0$ as a number is judged by him as an inappropriate grammatical inclination. His attitude here shows that he did not consider the criterion for acceptance or rejection of novelty a simple democratic process of establishing the major trends in the various communities of mathematicians. What distinguishes him from the conventionalist is the fact that, for him, the criterion is the legitimacy of subtle changes in the meaning of central symbols used. It is a grammatical mistake to say that 'has $\aleph_0$ elements', defined as 'can be 1-1 correlated with the set of naturals', is equivalent to the ordinary phrase 'has *p* elements' (Frascolla 1994, chap 3). This non-equivalence itself however is a matter of semantics, and thus embedded on the practice of the community.

This mathematical example illustrates some of the possible ramifications of the claim that the ultimate source of normativity in rule-following is located in the community as it engages in a given practice. Adopting this view of rule-following makes it easy to understand how people can make mistakes in trying to follow a rule. I may be mistaken in thinking that the exemplary cases should lead me in a particular direction, which then turns out to be at variance with that expressed in the practice in which I am engaged. For example, in linguistic practice, I may think that I fully know what the word 'cool' means, and then realise that it's use is also allowed in ways I didn't know of before. Ultimately, one realises that dictionaries and manuals of grammar do not occupy a definite, normative position over and above the linguistic practice, but the other way round. Hence, rule-following as here understood involves the negotiation with others so as to ascertain that the rule-follower is not mistaken.

This view has been challenged by those who argue that the source of normativity lies essentially within the individual (Blackburn 1984). Put simply, the argument starts from the observation that, when I am considering the third step in rule-following, in other words when I want to account for the possibility not only of extrapolating from the exemplary cases but also of getting it right or wrong, then I

need a guiding criterion to show me a red light if I go wrong. This criterion need not be something outside me, as was assumed up to now. It could be my disposition itself. Hence, I may negotiate with myself over time checking my present act of rule-following with previous instances when I followed the same rule. This checking may result in the discovery that I'm not really faithful to my disposition. In this case, my rule-following is wrong, and needs correction. Such an account of rule-following is especially plausible when dealing with regularities a person engages in, and in which nobody else does, for example when a person acts in a certain regular way to ease a certain kind of pain he cannot describe in words.

This account can be called an individualist view, while the previous one a communitarian view. Taken on their own, both seem necessary but insufficient. A combination of both is more plausible. This is what has been suggested by Philip Pettit (1993 p. 76-106). According to a combined view, rule-following is best seen in the following terms. First, I take certain cases to exemplify a particular rule; second, I extrapolate from the rule towards new cases by letting the disposition that has been thereby generated lead me on; third, I check my move. The guiding criterion here involves two aspects. First there is negotiation with myself over time, to check whether I am indeed faithful to my disposition. Then there is also negotiation with others to check whether the practice in which I am engaged confirms my action or not. It may happen that I find my inclination leading me to extrapolate in one direction while the inclination of my observer leads him repeatedly to extrapolate in another direction. In such cases, since there is no convergence between me and my observer, the rule I thought I was following does not exist. My project of following a rule would be empty because it lacks a target. Both self-consistency and consistency within the group are necessary conditions.

The hidden assumption here is that rules exist independently of their followers. This is certainly plausible. The rule, because it is being followed within a group, is, for that very reason, not completely dependent on any one single individual. To have a criterion of correctness, which is an essential feature of rule-following, one must have a certain distance between the rule and the rule-follower. Otherwise, one would end up with an impossible scenario, wherein, for example, all individuals just decide as they go along on how to use words. The obvious result would be a *dialogue de sourds*.

This account of rule-following, therefore, satisfies the two conditions for realism. The first condition was that the subject matter under discussion allows a mind-independent world. In this case, the rule itself is not dependent on the individual. The individual needs to negotiate with the community so as to ascertain that things are working well. The second condition for realism was that we could come to know something about the mind-independent world. On this account of rule-following, this is satisfied in the sense that the individual has a way of correctly following the rule determined by the community and thus a way of knowing ever more clearly what it is.

On closer scrutiny, however, a problem arises with the first condition. Is it enough to say that since the rule is independent of the individual rule-follower, then it is mind-independent? It seems not. What we have in this account is a situation where the rule depends on the community as a whole. This covers the case of *conventions*: it covers rules like 'Driving on the left side of the road is wrong'. But this is not enough for a robust sense of realism. It is not enough because the rule is not independent of the community as a whole. If, as regards criteria of correctness, we limit ourselves exclusively to the community as a whole, we would be in a situation similar to a choir with neither conductor nor score. The only criterion at work is for members of the choir to try to silence those singers whose notes are discordant with the harmony being attempted by the majority. This situation safeguards the idea of correctness as applied to the individual. What about the idea of correctness as applied to the whole group?

To arrive at a more robust realism as regards rule-following, one must consider how the community is constrained by states of affairs that go beyond the community itself. In other words, one must consider how the community is constrained by the nature of things, constrained by how the world is. Consider the set of all possible rules that a community can ever conceivably follow. We may call this a nomological space. A first major division of this space can be determined in terms of what is allowed by the physical laws of Nature and what isn't. Consider for example the following rule: 'No one can remain suspended in mid-air for more than 3 minutes'. This rule is definitely useless. Not that it doesn't make logical sense. It is useless, or impossible, in the sense that it cannot be a rule in the kind of universe human rule-followers live in. Any practice of any group is in fact limited by, or contained within, non-conventional constraints. So we can forget about the naturally impossible rules, and concentrate on the rest of the nomological space, containing naturally possible

rules. This set can be further subdivided into two groups: the conventionally possible rules and conventionally actual rules. Hence we may differentiate between, say, the rule 'one should drive on the left' from the corresponding rule 'one should drive on the right'. Both are naturally possible, but one is conventionally actual, the other not.

Further insight into this stronger realist account of rule-following may be gained by asking what can indeed act as a *non-conventional* constraint on practice. The first obvious set of constraints, as already mentioned, is determined by the boundaries of the physical laws of nature. The nature of things, therefore, becomes significant, at least in a negative way. Inasmuch as human rule-followers are physical things in the physical universe we know, they must act within these limits.

Human rule-followers, however, are not just inanimate objects. They are also biological organisms. There are laws governing biological organisms over and above the laws of physics. Hence, biological systems, apart from being under the common laws of physics, show additional characteristics, for example the tendency to grow towards a fixed structure, the tendency to reproduce, and so on. Human rule-followers are biological organisms. Therefore, any practice they engage in must be constrained not only by the laws of physics, that constrain all things, but also by the more specific characteristics of biological organisms. For example, as animals are characterised by the tendency to care for their offspring, a human rule that goes against this tendency would be a rule outside the legitimate constraints for human practice. It should be noticed here that these constraints are non-conventional. On this view, a rule that goes against the care of human offspring should be considered bad or senseless not because of established trends within the community but because it would effectively be neglecting the fact that humans are not just things, but special kinds of things, namely living things.

The final step may be taken by bringing in the further specification concerning humans. A human rule-follower is not only a thing constrained by the laws of physics, not only a biological organism constrained by the laws of living systems, but also a specific biological system characterised by rationality. To be true to their own specification, humans must engage in practices that do not go against rationality. This is not to say that human practices cannot be neutral as regards reason. What the constraint is blocking are practices that are irrational: for example building society on principles of ethnic cleansing, or on the systematic propagation of lies. Such constraints are non-conventional.

Admittedly, one can notice a difference in the binding force of the constraints mentioned so far. The constraints due to the laws of physics are rigid, in the sense that no thing can break the laws of physics. The constraints due to the characteristics of biological organisms are less rigid, in the sense that although all higher animals naturally care for their offspring, it is not inconceivable that one particular individual of the species neglects the constraint and still survives. The constraints due to the characteristics of human beings are even less rigid, in the sense that although all humans naturally act in accordance to reason, it is not inconceivable – in fact it is not rare – that one particular human person neglects the constraint by acting irrationally. In Aristotelian-Thomistic vocabulary, one would say that, as regards the practical judgement, because the factors in individual cases are indeterminably variable, a judgement concerning individual cases cannot be deduced with mathematical certainty but must be left to the *prudentia* of each person. The upshot here is that robust realism as regards rule-following involves referring to constraints that go beyond the community as a whole.[1]

My general aim in this paper was not to argue explicitly for realism and against antirealism as regards rule-following. It was to determine more clearly what realists argue for if they want to argue for their position. This was done by showing what realism involves when applied to debates on rule-following. The account presented above shows that realism in rule-following involves not only the weak claim that rules to be followed are independent of the *individual* rule-follower, as conventions are. It involves also the stronger claim that conventional rules are constrained by non-conventional constraints. These constraints depend neither on the individual nor on the group. They arise from the nature of things. Human rule-followers are constrained first by being material things in the universe that we know. They are constrained secondly by being living organisms. They are constrained as

---

[1] The account given here of non-conventional constraints on practice has been inspired by work done in the area of Natural Law, especially Thomas Aquinas, *Summa Theologiae*, I-II, Question 94, article 2. As regards the inevitable reference to *prudentia* (or *φρονησις*) for individual cases, the main reason seems to be the fact that we cannot, without falling into problems of infinite regress, conceive of rules for the correct application of rules. Aristotle was clear that 'the whole account of matters of conduct must be given in outline and not precisely, as we said at the beginning that the accounts we demand must be in accordance with the subject-matter; matters concerned with conduct and questions of what is good for us have no fixity, any more than matters of health. The general account being of this nature, the account of particular cases is yet more lacking in exactness.' (*Nicomachean Ethics*, Bk II, Chap 2.). In moral matters, our apprehension does not work in the style of Cartesian-deduction according to which we first fix the first principles and then deduce the particular case. See: Thomas Aquinas, *Commentary on the Ethics* I, lect. 12; II, lect. 2.

well by being rational. For a given practice, for example the practice of fishing or of celebrating birthdays, one may speak of the available space for possible conventional variety. From a realist point of view, this available space is limited. One cannot go fishing, or celebrate birthdays, in ways that go against the laws of physics, in ways that go against the general characteristics of biological organisms, in ways that go against right reason.

Two points of clarification are in order here. The first concerns the naturalistic fallacy, which may creep into the argument if one is not careful. If one accepts the account presented here, one may conceive the possibility that, in certain areas of human living, the available space for conventional variety, is discovered to be very restricted. In these areas, what we ought to do will be less dependent on community conventions and more dependent on how things are. This is not to say that, in these cases, the 'ought' is being *derived* from the 'is'. It means that knowledge of how things are is a presupposition of what ought to be done. The more we know how things are, including how we humans are, the better prepared we'll be to determine correctly what we ought to do. Once we determine, in a certain domain and to a certain extent, what we ought to do, we'll have a clearer idea of what being rational entails. [2]

The second point of clarification regards the meaning of the term 'rational' which should be well distinguished from ratiocinative. Rationality is not being taken to mean a system of rules, perhaps still unknown in their totality, common to all human beings, as Aristotle is sometimes interpreted to have implied when using the term rational to distinguish humans from other animals. Rationality is here taken to mean the ability to arrive at intelligent judgement that goes beyond ratiocination, that cannot therefore be captured in a system of rules. One task in fact of rational assessment is to determine which rules much be followed in a given situation. The view I'm taking here is that even though there may be some logical rules that form a common element for all humans, one should allow that humans have cognitive limitations and can therefore be in a genuine situation of rational disagreement, which can only be resolved after the various philosophical traditions involved in the debate

[2] John Finnis in his interpretation of Aquinas develops this point much further, not however in terms of rule-following. Cf. *Natural Law and Natural Rights*, Oxford University Press, 1980.

are well understood and appreciated by all.[3] Learning more about human rationality will enable us to determine more easily what we ought to do, but this learning procedure is a long communitarian endeavour.

Understood in these terms, realism in rule-following, because it allows a healthy symbiosis of conventional and non-conventional normativity, appears as an attractive position to defend.[4]

REFERENCES

Blackburn, S., (1984), 'The Individual strikes back', *Synthese* 58, pp. 281-301.

S. Kripke, *Wittgenstein on Rules and Private Language*, Oxford: Basil Blackwell, 1982, p. 21.

P. Boghossian, 'The Rule-Following Considerations', *Mind* 98 (1989), pp. 504-550;

G.P. Baker and P. M. S. Hacker, *Wittgenstein: Rules, Grammar and Necessity. An analytical Commentary on the Philosophical Investigations, volume 2*, Oxford: Blackwell, 1985;

S. Holtzman and C. Leich, (eds.), *Wittgenstein To Follow a Rule*, London: Routledge and Kegan Paul, 1981;

Pettit, P., 'The Reality of Rule-Following', *Mind* 99 (1990), pp. 1-21;

——, (1993), *The Common Mind*, Oxford: Oxford University Press.

C. Wright, (ed.), *Special Issue on Rule-Following*, *Synthese* 58 (1984).

Wittgenstein, L., (1953), *Philosophical Investigations*, trans. G. E. M. Anscombe, Oxford: Blackwell.

—— (1979a), *Remarks on the Foundations of Mathematics*, ed. G.H. von Wright, R. Rhees and G.E.M. Anscombe, trans. G.E.M. Anscombe, 3rd edn, Oxford: Blackwell.

—— (1979b), *Wittgenstein's Lectures, Cambridge 1932-1935*, A. Ambrose (ed.), Oxford: Blackwell.

---

[3] C.f.: Alasdair MacIntyre, *Whose Justice? Whose Rationality?* Notre Dame, Indiana: University of Notre Dame Press, 1988.

[4] Acknowledgements: A previous draft of this paper was given at the annual conference of the Società Italiana della Filosofia Analitica, entitled *Normativity, Fact, and Value*, held in Rome, 28 - 30 October 1999. Special thanks are due to participants who offered comments during the conference, and also to Prof. Philip Pettit, Research School of Social Sciences, Australian National University, Canberra.

P. Frascolla, *Wittgenstein's Philosophy of Mathematics*, London: Routledge, 1994, chapter 3.