

## Mindshaping and rules

Jaroslav Peregrin

intended for T. Zawidski, ed., *Routledge Handbook to Mindshaping*, Routledge

*If we look at mankind inter-culturally, we find an amazing diversity. If we look at mankind within any one community/culture, we find an equally amazing discipline and restraint. Question: how can a species, genetically granted by Nature such remarkable freedom and licence, nevertheless observe such restraint, such narrowly defined limits, in its actual conduct? Man is born genetically free but is everywhere in cultural chains.*

Ernest Gellner

**Abstract.** Our behavior is incomparably more flexible than that of any other animal species. The advantage is that we have become capable of surviving in different kinds of environments and withstanding even their abrupt changes; the disadvantage is that our behavior is *too much* flexible, in that at every moment we can do not only useful and reasonable things, but also things that useless and silly. We have solved this situation so that we complemented our "hardware" (genetic) freedom by a "software" (cultural) order. I suggest that the mechanism is based on rules; on our normative attitudes that we learn to assume in interdependence with learning to respect them. A vital part of the process can be seen as mindshaping, as learning to convey one's thoughts and actions into socially sanctioned channels. It is not only that the social cognition stores hard-won capabilities, which it instills on its newbies, it is also that it formats the cognitions of the individual members so as to chime with cognitions of other members.

### Our behavioral plasticity

Let me, to begin with, consider an old and almost forgotten paper by Ernest Gellner (1989). "What the human species does share genetically," writes Gellner, "is an unbelievable degree of behavioural plasticity or volatility". It is not so difficult to notice this, but Gellner urged that this is something characteristic of our species. He remarks that the actual differences among

people of different cultures cannot be a matter of genetics, but "what does seem genetically based in humankind is the plasticity, the volatility itself. All members of the species are endowed with it, and no other species possesses it." Gellner poses the following question: "How can a species, genetically granted by Nature such remarkable freedom and license, nevertheless observe such restraint, such narrowly defined limits, in its actual conduct?"

I think this is a very interesting viewpoint on humankind; and I think Gellner's is a very important question. And also I think that by now we can, at least roughly, answer it. I am convinced that the fact that our behavior is incomparably more flexible than that of any of our animal cousins has clear advantages (that are most probably behind it having been selected for during evolution). The advantage is that we have become capable of surviving in different kinds of environments and withstanding even their abrupt changes. But there are also disadvantages: our repertoire of behavior is so vast that at every moment we can do an awful lots of things: from the useful and reasonable ones to ones that useless and silly. From this backside, our behavior is *too much* flexible.

I will argue that we have solved this situation so that we complemented our "hardware" (genetic) freedom by a "software" (cultural) order. The reason is to be sought precisely in the fact that our behavior is *too much* flexible<sup>1</sup>. There are lots of tried and true ways of how to deal with the world, and there are a lot of sensible ways to deal with each other. But it would be a disaster if everybody were to seek these ways anew, by trial and error. If the environment changes, it may make sense to seek new ways; but unless this happens, it is much more reasonable to stay with the tried ones.

How does culture manage to preserve and replicate the useful ways of behavior? I think that the short answer is in terms of *rules*. Elsewhere I have discussed the nature of rules at length (Peregrin, 2014; 2024). An important point is that rules are not just explicit instructions - there exist rule-governed practices the rules of which are merely implicit to the practitioners' actions. However, this brings forth an important question: what is an implicit rule?

---

<sup>1</sup> The comparison of our nature and culture with hardware and software of a computer may be deeper than meets the eye. I contend that our, humans', becoming so behaviorally versatile is a direct parallel of computers' becoming universal. The thing is that universal computers became extremely powerful, because they could be turned, by means of software, into all kinds of special-purpose machines (like calculators, word processors, electronic diaries, game consoles etc.). However, without the software they would be just powerless. Cf. Rorty (2004).

## Implicit rules

When Wittgenstein (1953, §75) contemplated the nature of rules and of rule following, he came to face the question that not every rule could be based on an interpretation. Imagine that we see a signpost, which tells us that following red markings we reach a town in an hour. To understand this, we must understand what is written on the signpost, and indeed where does the signpost point. (Wittgenstein points out that even the latter is not self-explanatory – to understand it may require some enculturation.)

But if I must interpret the sign, it goes without saying that I must interpret it *correctly*. Hence I must follow some rules of the interpretation. If this rule following were to be again a matter of interpretation, we would need another rule and the infinite regress would be imminent. To avoid it, Wittgenstein thinks that we need some rules that are not explicit (in the form of signposts, linguistic instructions or whatever), the following of which does not involve interpretation. "Any interpretation still hangs in the air along with what it interprets, and cannot give it any support." (Wittgenstein, 1953, §198)

I propose a specific explanation of the nature of implicit rules (Peregrin, 2014; 2024). Inspired by Brandom (1994)<sup>2</sup>, I maintain that such rules exist in terms of *normative attitudes*, a specific practical pro- and con- attitudes of the members of the society in question to the ways they behave. It is important that the attitudes target only the *kinds* of behavior, not its protagonists. Thus if I oppose beating and support helping, I must oppose/support it independently of how it concerns me or any other specific person – whether it is me who is helped, or whether it is, say, Bob, who beats somebody<sup>3</sup>.

If the attitudes to a kind of behavior align across a society, we have a rudimentary implicit rule. Thus, e.g., if plus/minus all the members of the society tend to oppose beating (not just beating specific persons or by specific brawlers, but beating as such), there is a rule that beating is wrong, that one should not beat anybody. (Such a rudimentary rule will probably tend to evolve into something socially much more complex, perhaps with specific group of persons looking for violations of the rule and punishing it etc.)

A consequence of the fact that rules can exist via normative attitudes is that learning to follow a rule may be more of acquiring a skill, a know-how than acquiring an explicit knowledge, a know-that. We learn to recognize and avoid the "social friction" that our actions sometimes

---

<sup>2</sup> Due to the creative way I put Brandom's ideas to use, I do not want to make him responsible for anything I say, but I must acknowledge my deep indebtedness to him.

<sup>3</sup> But of course it may concern specific roles: a rule may be, e.g., that *nobody may beat anybody, with the exception of the chief, who may beat anybody* or that *nobody may beat anybody who does not deserve punishment* etc.

elicit and that is a matter of others' negative normative attitudes; and we learn to act so that our actions elicit positive normative attitudes. Especially, then, we learn to live within our normative spaces by palpating the limits the rules pose to us and learning to deal with them as borders.

### **The role of normative attitudes**

Imagine a driver weaving her way through a city duly observing all the rules of traffic. Imagine, then, a dog running through the same city and by chance doing this also in accordance with traffic rules. It would be hard to deny that there is a basic difference between these two cases. While the driver follows the rules of traffic in the sense of doing this intentionally, the dog just happens to comply with these rules without knowing it. The way we have just formulated the difference seems to indicate that it is located in the minds of the protagonists. While the driver has the rules, as it were, before her mind's eye, nothing such can be said of the dog. However, what exactly is it that is present in the mind of the human and absent from the mind of the dog?

It cannot be linguistic articulations of the rules: one can learn to follow the rules of traffic without ever verbalizing them, perhaps merely by trial and error (if she is lucky enough to survive the errors). It cannot be the knowledge of what to do in every possible situation: there is an infinite number of such situations. And in general, there doesn't seem to be a suitable mental object whose presence in the mind would render a behavior rule-governed and whose absence would signal the absence of the rule (Kripke, 1982).

The question that is being answered here is generally the following: how does a behavior governed by a rule differ from other kinds of behavior, especially from a behavior that is merely regular? And the answer I suggest is that we must seek the difference not somewhere "behind" or "beyond" the behavioral pattern we are investigating (*viz.* in the minds of its protagonists), but rather "besides" the pattern, in a broader context of the occurrence of this pattern. If an individual not only stops at red traffic lights, but also is angry with those who do not, if she not only observes speed limits, but also dutifully pays a fine in case she happens to exceed them, and if she not only gives way, at crossroads, to those who have priority, but also counts on others doing the same, then she - unlike the dog - can be seen as following the rules of traffic.

It follows that without a required context there may be sometimes no answer to the question whether somebody does, or does not follow a rule. Does it mean that in such a case there is no difference at all? Not necessarily. There might be a difference which could be perhaps diagnosed by probing one's brain: the driver might, for example, think about the rules and

control her driving observing them (which might be, in principle, revealed by a brain scan). But it is also possible that she drives absentmindedly, without even realizing that there is something as rules of traffic, and in such a cases the answer might be unavailable.

Does it mean that rule following has nothing to do with minds? Not really. The point is that genuine rule following presupposes normative attitudes and assuming a normative attitude is not an occurrent event in the mind, but rather a relatively stable state. We do not stop to assume the attitude when we are not thinking about it, and not even when we are sleeping. Assuming such an attitude is something like supporting a sport team. If somebody goes to see a match of the team, it is an indication of the support, but we cannot be sure that he really supports it. We need a broader context.

Of course that it is possible to follow a rule without assuming the corresponding normative attitudes - one may be forced to follow it without accepting it. On the other hand, a rule, especially an implicit rule, cannot exist without the support of corresponding normative attitudes. If the rule is not to fade away, there must be a monitoring of its violations and a mechanism of rectification of possible deviations. This is a direction that the concept of normative attitudes leads us: a behavior is rule-governed only if it is generally measured by what *ought to be*, and it is so measured by the *behavior* of the protagonists of the allegedly rule-governed practice.

## **Social rules**

There are basically two kinds of rules that are preserved (and perhaps upgraded) culturally. First of them are what von Wright (1963) calls *technical norms* or *directives*. They are *instrumental* rules, rules which help us achieve something of an enduring value for us. Typically, they concern our dealing with nature: it may be rules of how to build a kayak, how to hunt down a deer, or how to construct a nuclear reactor. A lot of recent anthropologists put stress on culture as a storage of such instructions (Boyd & Richerson, 2005; Herich, 2015). But there is another important kind of rules, which are usually called *social*. These rules have to do with our cooperation and with our peaceful and productive coexistence with each other. These rules have to do not with effectiveness, but mostly with predictability, for predictability is the mother of peaceful coexistence. Mercier & Sperber (2017) argue that here is also the root of our all-important “game of giving and asking for reasons” and consequently of our reasoning – that giving reasons took origin from the situation when one does something unpredictable and wants to rectify this.

It may seem that just as the first kind of rules we have mentioned help us operate successfully within the realm of nature, this other kind of rules helps us operate successfully within the

realm of our society. But this is not the whole truth. True, we *may* use social rules as directives to help us get what we want from our peers just like we use the instrumental rules to help us get what we want from the natural world. These rules *may* help us fruitfully navigate the landscape of the social world like the others help us navigate that of the natural one. However, what we call social rules also *creates* the landscape of the social world.

The natural world is not easy to deal with because it puts up resistance: many things we need or want is not easily got from it and we must invent clever ways to get them. Also the social world wields resistance - but here the most intensive and the most systematic part of the resistance is normative - it is a matter of implicit rules carried by our attitudes, as well as of explicit rules, embodied in our codes and rule books.

Imagine a rule that that says that older people should be greeted. It can be used as a directive: you may, e.g., derive that if you want something from your peers, it is good to greet at least those of them who are older. However, why does this work? Why is this a good course to steer the social landscape? Because if you do not follow it, you may encounter a resistance: you may be scorned by your peers and they may be less forthcoming to you. And why is that? Because they all accept the rule in the sense of evincing the corresponding normative attitudes.

Imagine, in contrast to this, the directive that you should not step on a thin ice. Why is this a good course to steer the natural landscape? Because if you do not follow it, you may encounter a resistance: the ice may break and you may drown. This danger is independent of what other people do, it is a matter of nature alone.

Hence while a prototypical directive tells you what you should do (to get what you want) to overcome an objective resistance, in case of social rules, the resistance that is to be overcome, is of your making (of course not of you alone, but of you and your peers). When you learn to obey the rules, sooner or later you understand that you, as everybody else, is to not only to submit to the rules, but also come to take part in their enforcement, i.e. evince the normative attitudes that keep them in being<sup>4</sup>.

### **Mindshaping**

Zawidzki (2013) argues that that an important part of human evolution was the development of the capacity which he called mindshaping, i.e. "making human minds and behavior more homogeneous and hence easier to predict and interpret" (p. 29), This, according to him, must

---

<sup>4</sup> Cf. Peregrin (forthcoming-a).

have preceded the broadly discussed phenomenon of mindreading (Nichols & Stich, 2003; Lurz 2011), *viz.* the capacity of estimating the intentions of one's conspecifics.

Elsewhere (Peregrin, 2020) I pointed out that the crucial role within the process of mindshaping must have been played by our human invention of *rules* and here I am building on this finding. We have seen that humans have compensated for too much volatility of their behavior by the establishment of publicly sanctioned "channels" into which they convey each other's behavior. And it is the normative order of our social world which does this work.

Hence out of the vast number of things we can do, we do only some distinguished ones, or we do things in distinguished ways. However, it is not so that we *decide* for the socially accepted courses of action, the straight jacket of our society becomes our "second nature" to such an extent that we just act in the expected ways and deviate from them only for substantive reasons. This is the result of the fact that the normative attitudes, which permeate the whole society, act like walls: they reliably deflect us from certain courses of action letting us move only in certain directions.

Now what holds for our behavior holds for our thinking: it too was conveyed to socially hollowed channels. This may be hard to understand for those who are convinced that mind causes behavior, but behavior does not influence mind; but I do not think this is correct. I side with Tomasello (2014, p. 38):

The general process is thus that the young child imagines how some social interactant is comprehending or evaluating her, and then she uses this to socially self-regulate. Scaling up the sociality involved, children from about three years of age (but, needless to say, not apes) socially self-regulate on the basis of cultural structures— such as, prototypically, conventional and moral norms— that are based in cognitive processes of collective intentionality, what we may call normative self-governance. (...) Thus, from sometime during the late preschool period, young children self-regulate both their thinking and actions not just by how efficacious they will be in the current context (as do apes), and not just by how they will affect a particular person's thoughts or evaluations (as do younger children), but also by the perspective of how these will fit with the normative expectations of the social group. This process essentially constitutes the construction of a normative point of view as a self-regulating mechanism, arguably the capstone of the ontogeny of uniquely human cognition (normative rationality) and sociality (normative morality).

Thus I think that putting up the normative infrastructure of a society we not only regulate behavior, but also thinking; or maybe we make people self-regulate.

### **Soft- instead of hard-wiring**

Anyway, it would seem that rules let us retain a "soft" way of keeping doing the "right" things. (Here "right" might mean tried and true, or conforming to social conventions.) In simpler organisms, such ways may be "hardwired": the organism is simply set up so as to do things in these ways come what may. But we have seen that in our case this would be a hindrance - the trajectory throughout evolution to which we converged rests on flexibility. Hence we cannot make do with the "hardwiring"; but as abandoning it without a replacement would be equally disastrous as having it, we have developed the soft version<sup>5</sup>.

Compare this with our emotions, the status of which is similar. Simple organisms may have hardwired connections between perception and behavior. In our case (and in case of many other complex organisms) this would not work: we need a more flexible management of behavior. (Imagine that we are hungry and see a nice apple on a nearby tree. If we were determined to follow our hunger and to go for the apple come what may, we would not be able to avoid a predator which may turn out to be hidden in the tree. We need a flexible enough management of behavior that allows us to run away, letting the apple be.) The solution are emotions: they guide our behavior along the tried paths, but not in a rigid way. A more pressing emotion can trump a less pressing one, allowing the relatively optimal behavior result from the interplay.

However, more complex patterns of behavior cannot be produced by emotions alone. Here what is employed (over and above emotions) is a "collective memory" of the relevant community - the patterns are kept in place via rules, which are in turn upheld by normative attitudes. This is a unique solution which has brought us, humans, to our incredibly complex culture<sup>6</sup>.

Not that normative attitudes, rules and culture would work independently of emotions; indeed, rules often build on the infrastructure of emotions. But the emotions are often not so determinate to route them in a specific direction. Emotions like pride or shame must be filled with a specific content, which is what rules and culture is capable of doing. Hence such

---

<sup>5</sup> Zawidzki (2018) also uses the metaphor of hardware/software. He points out that software, i.e. programs written in programming language, does not consist of descriptive statements, which aspire to represent the world (have the "language-to-world" direction of fit) but rather of commands, which attempt to make some changes in the world (have the "world-to-language" direction of fit), especially in the memory of the computer. Similarly, the "social software" is here not to equip individuals with tools to capture their inner world, but rather with tools that give the inner world its shape.

<sup>6</sup> See Peregrin (2024).



emotions can be seen as parts of a do-it-yourself kit which culture can use to build the normative network innervating a society.

It even seems that from this vantage point, mind developed precisely as a thing that can be soft-wired (or programmed, if you want), once the hard-wiring turns out to be insufficient. Mind is the medium where you can imprint various things, including the optimal paths your behavior is to respect, without this being forever and unchangeable - mind can be updated.

But note also that mind alone is insufficient - a crucial work must be done by a *community* of minds. We need that the optimal paths get imprinted into every new mind, and this necessitates a mechanism for the imprinting, viz. culture. Therefore, an individual mind must be equipped not merely by the imprint, but by the normative attitudes that effect it and that can be passed along from generation to generation.

Zawidzki (2021) entertained similar ideas in the context of the relationship between individual metacognition and social cognition - his idea is that social cognition provides roles and categories which individuals employ for their metacognition, thus using the metacognition not for mapping an already existing cognition, but rather for giving it the socially constituted shape. Zawidzki writes:

Person-level, linguistically expressible, metacognitive concepts are socio-cognitive tools that individuals acquire from their cultures, and that transform them as cognitive agents and cognitive objects, making them better at predicting others and easier to predict by others. Thus, they have both regulative and descriptive/predictive functions, but the latter depend on the former: we can use our metacognitive concepts to describe and predict each other's behavior because we have used them to regulate our own and each other's behavior in ways that make it easily describable and predictable in their terms.

Elsewhere (Peregrin, forthcoming-b) I suggested that human life form is characterized by being structured into practices that are "specified by a system of rules which defines offices, roles, moves, penalties, defenses, and so on, and which gives the activity its structure" (Rawls, 1955, p. 33). I tried to anatomize the way in which such practices (and institutions) are constituted by rules and came to the metaphor of virtual arenas constituted by systems of rules, in which we have come to dwell and which let us carry out actions which are not available to us outside of the arenas.

Being born into a human society, we enter, during an early stage of our upbringing, the most important of such arenas (like the arena of morality and that of language) and learn to live in them. Also we learn how to visit many other arenas (e.g. those of various games); and later we learn how to help maintain the arenas and possibly how to help create new ones. The

arenas let us live our “unnatural” lives (as instead of within the milieu of nature we largely live them within the “virtual” milieus of our own making). And this whole labyrinth of normative spaces is made possible by that we have submitted ourselves to the discipline of rules and have our minds (benignly) shaped.

## **Conclusion**

Genetic hardwiring of behavioral patterns is a mighty tool; however, from some point of complexity of organism it can become a hindrance. More flexible solution is needed; and it is in our case, the soft-wiring of the patterns by means of structures preserved via the social cognition. I suggest that the mechanism in which such enculturation works is based on rules; on our normative attitudes that we learn to assume in interdependence with learning to respect them. I also suggest that a vital part of the process can be seen as mindshaping, as learning to convey one's thoughts and actions into socially sanctioned channels.

We humans are social animals; and we have driven our sociality to an unprecedented form of cooperation and mutual dependence. This, on the one hand, is a matter of course, while, on the other hand, it is often not thought through to its consequences. In particular, there is a unique interplay between social cognition and individual cognitions of members of the society. It is not only that the social cognition stores hard-won capabilities, which it instills on its newbies, it is also that it formats the cognitions of the individual members so as to chime with cognitions of other members.

Gellner's paper that I quoted at the beginning of this article closes in the following way (p. 525):

My argument has been that genetic under-programming must have been linked to the presence of a compensating system of cultural/linguistic restriction. These cultural systems, and systems of coercion, have complemented each other in diverse ways at different stages. The volatility must obviously have had its own genetic preconditions, so that our volatility, and our endowment with compensating talents and propensities, must have arrived jointly. The consequence has been the emergence of a species in whose life both social or semantic transmission and institutionalized coercion have become far more important than genetic mutation, making it possible for change to be astonishingly rapid.

I subscribe to these words of my virtual compatriot<sup>7</sup>.

## References

- Boyd, R. & Richerson, P. J. (2005): *Not by genes alone*, Chicago: University of Chicago Press.
- Brandom, R. (1994): *Making it explicit: Reasoning, representing, and discursive commitment*, Cambridge (Mass.): Harvard University Press.
- Gellner, E. (1989): 'Culture, constraint and community: semantic and coercive compensations for the genetic under-determination of Homo sapiens sapiens'. *The Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans*. Edinburgh University Press, Edinburgh.
- Henrich, J. (2015): *The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter*, Princeton: Princeton University Press.
- Kripke, S. A. (1982): *Wittgenstein on rules and private language: An elementary exposition*, Cambridge (Mass.): Harvard University Press.
- Lurz, R. W., ed. (2011): *Mindreading Animals: The Debate over What Animals Know about their Minds*, Cambridge (Mass.): The MIT Press.
- Mercier, H. & Sperber, D. (2017): *The enigma of reason*, Cambridge (Mass.): Harvard University Press.
- Nichols, S. & Stich, S. P. (2003): *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*, Oxford: Oxford University Press.
- Peregrin, J. (2014): *Inferentialism: why rules matter*, Basingstoke: Palgrave.
- Peregrin, J. (2020): 'Normative Mindshaping and the Normative Niche'. *Groups, Norms and Practices* (ed. L. Koreň et al.), pp. 85–98, Cham: Springer.
- Peregrin, J. (2024): *Normative Species: How Naturalized Inferentialism Explains Us*, New York: Routledge.
- Peregrin, J. (forthcoming-a): 'Normativity between philosophy and science'. *Philosophical Psychology*.
- Peregrin, J. (forthcoming-b): 'Inside Human Practices', W. Rzepiński et al. (eds.): *Practices in Legal and Social Theories*, New York: Routledge, forthcoming.

---

<sup>7</sup> Like me, Gellner spent the first part of his life in former Czechoslovakia (though I lived there significantly later than him). The onset of Nazism in Germany made his family leave for Britain.

- Rawls, J. (1955): 'Two Concepts of Rules'. *Philosophical Review* 64(1), 3–32.
- Rorty, R. (2004): 'The brain as hardware, culture as software: Symposium: Vincent Descombes, The Mind's Provisions'. *Inquiry* 47(3), 219–235.
- Von Wright, G. H. (1963): *Norm and Action*, New York: Humanities Press.
- Wittgenstein, L. (1953): *Philosophische Untersuchungen*, Oxford: Blackwell; English translation *Philosophical Investigations*, Oxford: Blackwell, 1953.
- Zawidzki, T. W. (2013): *Mindshaping: A new framework for understanding human social cognition*, Cambridge (Mass.): MIT Press.
- Zawidzki, T. W. (2018): 'Self-Interpretation as Software: Toward a New Understanding of Why False Self-Conceptions Persist'. *Third-Person Self-Knowledge, Self-Interpretation, and Narrative*, pp. 115–144, Springer.
- Zawidzki, T. W. (2021): 'A new perspective on the relationship between metacognition and social cognition: Metacognitive concepts as socio-cognitive tools'. *Synthese* 198(7), 6573–6596.