

**On ‘common-sense ontology’: a comment on the paper by Frank Hindriks and Francesco Guala**

**Robert Sugden**

School of Economics  
University of East Anglia  
Norwich NR4 7TJ  
UK

[r.sugden@uea.ac.uk](mailto:r.sugden@uea.ac.uk)

**Abstract** This note comments on Hindriks and Guala’s ‘unified theory of institutions’. One of the components that Hindriks and Guala seek to unify, and which they claim is unsatisfactory on its own, is the analysis of conventions that derives from the work of Lewis. I argue that the Lewisian approach provides simple and powerful explanations of many regularities in the social behaviour of humans and other animals. Those explanations can be seen as good social science even if, as Hindriks and Guala argue, they do not fit with common-sense ontology.

Hindriks and Guala (2014) propose a ‘unified theory of institutions’ which is intended to explain ‘not only what ordinary people talk about when they talk about institutions, but also how these objects relate to the theoretical constructs that social scientists use to explain coordinated behaviour in complex strategic interactions’ (pp. 19, 20). This theory unifies three different ‘conceptions of institutions’ that have been used by social scientists and philosophers, each of which (Hindriks and Guala claim) is unsatisfactory on its own. One of these conceptions, the ‘equilibrium-based conception’, is attributed to Lewis (1969), Ullman-Margalit (1977), me (Sugden, 1986), and to various later writers. I will argue that the approach that I have followed in explaining conventions, building on the work of Schelling (1960) and Maynard Smith (1982) as well as Lewis and Ullman-Margalit, is suitable for its intended purpose. It is perhaps not a ‘theory of institutions’ in the sense that Hindriks and Guala use this term, but I am not persuaded that constructing the kind of theory they have in mind is an appropriate objective for social science.

It seems that when Guala and Hindriks say that they are looking for a theory of institutions, they are using ‘institution’ to refer to whatever this word describes in ordinary language, rather than to any specific domain of observable behaviour: they want a theory that can connect with ‘common-sense ontology’ (p. 20). I cannot see why the objects of social-scientific explanation, any more than those of biology or chemistry, need to correspond with common-sense ontology. Expecting such a correspondence is particularly problematic when one is dealing with a concept – or perhaps a family of related concepts – identified by a word with a wide range of ordinary-language meanings. (The *Concise Oxford Dictionary* defines ‘institution’ as ‘established law, custom, or practice’.)

One should be especially wary of common-sense ontology when considering similarities between the behaviour of humans and other animals. Hindriks and Guala claim as a merit of their theory that it excludes ‘so-called animal conventions’ which ‘we would not intuitively consider institutions’ (p. 8). But is this really a merit? As Hume (1739–40/ 1978, p. 177) recognised when developing his seminal explanation of human conventions, there are many similarities of physiology and of behavioural and affective responses between humans and other intelligent animals. In Hume’s words, it would be contrary to ‘all our principles of reason and probability’ not to attribute like effects to like causes. Thus, in explaining phenomena that are common to human and non-human animals, there should be a scientific presumption in favour of theories that apply across species. In contrast, common-sense

ontology is littered with assumptions about human exceptionalism which philosophers are often reluctant to question. As Hume says:

The common defect of those systems, which philosophers have employ'd to account for the actions of the mind, is, that they suppose such a subtilty and refinement of thought, as not only exceeds the capacity of mere animals, but even of children and the common people in our own species; who are notwithstanding susceptible of the same emotions and affections as persons of the most accomplish'd genius and understanding.

My early work on conventions, which pre-dated the use of evolutionary game theory in economics, was inspired by Maynard Smith and his collaborators' investigations of animal conflicts. I was struck by the similarity between the way in which the asymmetry between 'possessor' and 'challenger' was used to resolve animal conflicts and the significance of the same asymmetry in resolving conflict in human life. I argued that these similarities reflected isomorphisms between natural selection and experiential learning (both of which are implicated in human behaviour and in the behaviour of other intelligent animals). My methodology, like that of Hume and Adam Smith (1759/ 1976), was to try to explain complex patterns in human social behaviour by using the simplest possible assumptions about individual psychology. On the principle of Ockham's razor, I took it to be a merit of a theory to be able to dispense with unnecessary assumptions.

My analysis rested on two properties that human reasoning is known to possess, and which are common to many other intelligent animals. The first is that, through experience and through positive and negative reinforcement, individuals learn to follow rules of behaviour that tend to be rewarded in the environment in which they are acting. Because no two decision problems are exactly alike, such rules are necessarily based on perceptions of similarity. Hindriks and Guala are of course right to say that human agents often represent these rules to themselves symbolically; but the importance of experiential learning in the animal world strongly suggests that such representation is not essential. The second psychological property is emotional contagion – the tendency for affective states, particularly fear and distress, to be transmitted from individual to individual (see, e.g., Preston and de Waal, 2002). In a setting in which individuals of the same species interact with one another, each individual's behaviour forms part of the environment to which other individuals' behaviour has to adapt. Thus, complex social regularities of behaviour and of affective repertoires can be emergent properties of repeated interaction, without any social-level reasoning. It is perhaps true (as Guala and Hindriks say) that speckled wood butterflies 'cannot invent a new equilibrium' (p. 9); but it seems to me that, in any interacting population

of an animal species that is capable of reinforcement learning, it must be possible for new equilibrium selection problems to be confronted and for new solutions to emerge. (The vocalisations of songbirds might be an example.) In other words, ‘social realities’ can emerge without their being (in Guala and Hendriks’s expression) ‘collectively accepted’.

I do not mean to deny that, by virtue of our powers of symbolic reasoning and our access to language, we human beings can do more than my analysis of conventions is designed to explain. Lewis (1969) shows how a superstructure of higher and higher tiers of ‘reason to believe’ can be grounded on simple, commonly-observed regularities of behaviour. Thus, for agents who are capable of sufficiently sophisticated reasoning, conventions are indeed collectively accepted. But the existence of the conventions themselves – that is, of regularities of behaviour and of affective response – need not depend on that sophisticated reasoning. (This interpretation of Lewis’s analysis is defended by Cubitt and Sugden, 2003.) Guala and Hendriks might perhaps reply that my definition of ‘convention’ is too permissive for their purposes – that it would let in too many behavioural patterns that common-sense ontology would not consider to be institutions. But ‘permissiveness’ is just another word for generality. I cannot see why the domain of a social scientific theory should be constrained to fit pre-scientific intuitions.

## References

- Cubitt, Robin and Robert Sugden (2003). Common knowledge, salience and convention: a reconstruction of David Lewis’ game theory. *Economics and Philosophy* 19: 175–210.
- Hendriks, Frank and Francesco Guala (2014). Institutions, rules, and equilibria: a unified theory. *Journal of Institutional Economics*, published online: 16 October 2014, DOI: 10.1017/S1744137414000496.
- Hume, David (1739–40/ 1978), *A Treatise of Human Nature*. Oxford: Clarendon Press.
- Lewis, David (1969), *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Maynard Smith, John (1982), *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Preston, Stephanie and Frans de Waal (2002). Empathy: its ultimate and proximate bases.  
*Behavioral and Brain Sciences* 25: 1–20.

Schelling, Thomas (1960), *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Smith, Adam (1759/1976). *The Theory of Moral Sentiments*. Oxford: Clarendon Press.

Sugden, Robert (1986), *The Economics of Rights, Co-operation and Welfare*, 2nd edn, 2004, Oxford: Blackwell.

Ullmann-Margalit, Edna (1977), *The Emergence of Norms*, Oxford: Clarendon Press.