

On defining institutions: rules *versus* equilibria

GEOFFREY M. HODGSON*

Hertfordshire Business School, University of Hertfordshire, Hatfield, Hertfordshire AL10 9AB, UK

Abstract. In their stimulating paper, Hindriks and Guala (2014) bridge the prominent alternative conceptions of institutions-as-rules and institutions-as-equilibria, by proposing a ‘rules in equilibrium’ interpretation. This comment argues that the task of *defining* institutions as a class of phenomena is different from the tasks of *understanding* or *analysing* them. Definitions are classification devices and are typically ill-based on behavioural outcomes such as equilibria. Accepting the useful insights of the Hindriks and Guala (2014) article, attention to the matter of definition reinstates a rules-based approach, notwithstanding the importance of understanding and analysing equilibria. The comment establishes a broad definition of institutions as systems of rules, which includes organizations. Finally this comment raised some of the problems involved in understanding the nature of institutional rules.

Science is a social process, wherein communication is central.¹ Notwithstanding the importance of mathematics and statistical data, much scientific communication will always involve ordinary language, with all its trappings and ambiguities. Some degree of clarity is required in the use of words. Accordingly, when an academic author uses the word ‘institution’, he or she should be able to point with adequate lucidity to the class of phenomena to which the term is attached. Absolute precision may be impossible, even if it were desirable. And definitions in the social sciences are likely to have fuzzy boundaries. But these are not excuses for abandoning the tasks of definition and the vital communication of meaning.

Definitions are not mere word-play. Comprehending the meaning of a word is often tied up with the understanding of real phenomena: the search for understanding drives scientific enquiry. Questions concerning the nature and meaning of gravity drove both the Newtonian and the Einsteinian Scientific Revolutions in physics. Scientists must first establish an agreed rough

*Email: g.m.hodgson@herts.ac.uk

¹ This comment uses some material from Hodgson (2015).

understanding of the phenomenon they are investigating. Then they try to focus on the problem, using a definition as a means of demarcation. They sometimes change this definition. Definitions matter at every stage. Science is driven in part by a search for meaning. Absolute precision, like absolute cleanliness, is impossible; but that does not imply that we should abandon our duties of linguistic housekeeping or personal hygiene.

We must seek some precision with words, to help us dissect an often messy reality with sharp concepts of the mind. A definition must identify essential features of the type it defines. It cannot simply be a description of the entity or group of entities. Definitions point to specific essential features: they are not mere lists of attributes.

Although, Hindriks and Guala (2014) make reference to some of the philosophical literature on definition, I suggest that their article is more concerned with understanding the general nature of institutions, rather than defining them. I propose that these two tasks are related but basically different. I wish to underline an established, basic definition of an institution in terms of social rules.

There are different types of definition (Robinson, 1950). Definitions in mathematics make clear-cut, axiomatic designations. Other kinds of definition are more concerned with the demarcation of one real-world species of entity from another, among an empirical ensemble of varied entities. This basic problem of empirical demarcation is found in both social sciences and biology, where there is typically substantial variation and a lack of pure types. In these disciplines, the role of a definition is to demarcate and assign a term to a type of entity – to distinguish one species from another, with possible fuzziness and awkward boundary cases. The role of such a definition is to identify the essential distinguishing characteristics, or to ‘carve’ reality ‘where the joint is’, as Plato reported of Socrates in *Phaedrus*.

When I was at school, I was taught that mammals are defined as a clade of animal where the females suckle their young. I also learned of strange examples, such as the duck-billed platypus, which suckles its young but also lays lizard-like eggs. I was informed that the suckling criterion is sufficient to demarcate the whole clade – including marsupials, mice, mammoths, walruses and whales – from other living things, including reptiles, birds and insects. One interesting feature of this suckling definition is that it does not attempt to describe mammals in more detail, or to list a host of important common features. It omits crucial descriptive characteristics, such as being warm-blooded vertebrates with four limbs, or having lungs and brains. The suckling criterion is sufficient to demarcate, and hence it can serve adequately as a definition. Definitions of this kind are neither analyses nor summary descriptions of classes of entities.

I have since discovered that the mammal clade is now defined in a different way, often involving jaws and teeth rather than the suckling of infants. The

leading palaeontologist Kemp (2005: 3) defined mammals as ‘synapsids that possess a dentary-squamosal jaw articulation and occlusion between upper and lower molars with a transverse component to the movement’.² In Kemp’s view, this criterion is sufficient to identify the clade originating with the last common ancestor of both *Sinoconodon* (an early type of mammal found as fossils in China) and more recent mammals. This shift of definition is tied to the need to classify fossils, where mammary glands and other soft tissue may not have survived. Specific features of jaws and teeth are sufficient to demarcate mammals from other synapsids, and to dispense with the suckling criterion.

What has this mammal story got to do with defining social phenomena such as institutions? We can learn from biology but social science encounters additional classificatory problems. In biology the formulation of definitional demarcation criteria is not easy, but it is facilitated by the fact that the evolution of species is hierarchical or tree-like: when a new species emerges it diverges from the others, and its members rarely if ever inter-breed across species boundaries.³ By contrast, in the social sciences, not only individuals but whole species can meld together, creating hybrid types as well as hybrid individuals. For example, the English language is a fusion of multiple language types, with Latin, Germanic and other separate roots. Some prominent legal systems involve combinations of common and civil law. Other institutions that develop in one country or setting are frequently copied and installed in another. Social evolution widely involves reticulate rather than hierarchic patterns of descent. Consequently, a specific species of social phenomenon cannot be conceptualized in terms of a single, earliest, common ancestor, from which all others in tree-like formation descend. As a result it is ineffective to identify a few key novel features shared by that earliest ancestor and its descendants, and then use these features as the definitional criteria.

Given such complications, where else can we turn for help? In his *Topica*, Aristotle wrote: ‘A definition is an account that signifies the essence’. Note that he used the verb ‘signify’, rather than ‘describe’ or ‘explain’. We need a sign: we have to find a common feature that signals the nature of the beasts, and it is sufficient to demarcate them from the others. Definitions do not need a fuller description. After two millennia of usage among scholars, Aristotle’s notion of essences has gone through post-modernist hard times, but arguably has survived the ordeal (Hodgson, 2015; Nussbaum, 1992; O’Neill, 1998; Sober, 1980). As O’Neill (1998: 8) put it: ‘The essential properties of an entity of a particular kind are those properties of the object that it must have if it is to be an object of that kind.’ Definitions identify the minimum number of essential

2 Compare with Rowe (1988).

3 By prominent accounts, the limits of a species are defined in terms of the possibility, or otherwise, of inter-breeding (de Queiroz, 2005).

properties that are sufficient to demarcate entities of that kind from other entities.⁴

As Aristotle explains in his *Metaphysics*, the essence of what an entity *is* cannot be adequately defined in terms of what an entity *does*, or by patterns that it generates. If we make this confusion, then we wrongly imply that when the entity interrupts its characteristic activity, then it ceases to be such an entity. Sometimes birds fly. But what defines a bird is the (existing or ancestral) *capacity* to fly, not flying itself. If a bird was defined as a flying animal, then any bird sitting on a branch or pecking on the ground would cease to be a bird. If a firm was defined as an organization producing goods or services, then when the workers were on holiday the firm would cease to exist. It would be better to define a firm as an organization with the capacity to produce goods or services. This may seem a small point, but the substitution of behaviours for dispositions or capacities is one of the major methodological errors in the social sciences.

An enduringly relevant point here is that definitions or ontologies that are based principally on behaviour cannot cope with instances where the behaviour radically changes or ceases. But the capacity to produce the original characteristic behaviour remains; this potentiality, not the outcome, constitutes the essence of the entity. Although ancient, this point is not arcane; it is widely utilized in modern realist philosophy of science (Bhaskar, 1975; Harré and Madden, 1975; Lawson, 1997; Popper, 1990). Central to most strands of modern realist philosophy is the distinction between dispositions and outcomes, where the former are more fundamental than the latter.

This leads to the question whether equilibria should be part of the definition of an institution. Being outcomes of individual interactions, equilibria are secondary to the relational framework that generates their possibility. In game theory, rules make up the *game form*, and an equilibrium is a possible *game outcome*.⁵ Social equilibria are important phenomena. But even if the equilibrium concept is vital to understand how institutions work, this does not mean that it has to be part of the definition. As Robinson (1950: 80) put it in his treatise on definitions: ‘The supreme rule of [definitional] stipulation is surely *to stipulate as little as possible*.’

There are many equilibria that do not themselves denote institutions. Physicists, biologists and chemists tell us that there are equilibria in nature, but these equilibria are not regarded as institutions. Consider market equilibria,

⁴ The word *essential* is ambiguous. It can refer to something that is part of the essence of a kind, or to something that is vital to its existence. Here *essential* refers to part of the essence of a type. Hence, for the purposes of definition, listing the ‘essential properties of an entity’ does not mean listing all the things that are vital for its survival.

⁵ Note that in pioneering uses of game theory to analyse institutions (Aoki, 2001; Binmore, 1994, 1998; Schotter, 1981), the resulting game equilibria themselves establish possible rules. For example, with the coordination game of driving-on-the-left *versus* driving-on-the-right, the two possible equilibria are alternative (regulative) rules.

denoted by stable prices and quantities. Markets may be institutions, but market prices are not. Furthermore, as 1929 and 2008 demonstrated, markets are also capable of highly disruptive disequilibria.

Crafters of definitions should also acknowledge the weight of customary meanings of words (Robinson, 1950: 79). The depiction of institutions as shared systems of rules has near-customary status in the literature on the topic (Crawford and Ostrom, 1995; Knight, 1992; Mantzavinos, 2001; North, 1990; Ostrom, 1990).

I propose elsewhere that the definition of institutions in terms of *integrated systems of rules that structure social interactions* (where the term ‘rule’ also requires definition – see below) gives us a rough but useful demarcation criterion to distinguish institutions from other social phenomena (Hodgson, 2006, 2015).

Typically, social theorists broadly define social structures as sets of relations between individuals (Porpora, 1987). Consequently, all institutions are social structures, but not all social structures are institutions. For example, the demographic or gender structure of a society is part of its social structure; but they are not essentially systems of rules, and hence they are not institutions. The ‘shared systems of rules’ definition of an institution does the job of excluding these important non-institutional phenomena: it is an effective demarcation device.

These rules include norms of behaviour and social conventions, as well as legal rules.⁶ Accordingly, systems of language, money, law, weights and measures, traffic conventions, table manners, and all organizations are institutions. But not all institutions are organizations. For example, language is an institution but not an organization.

Some have expressed uneasiness with this definition of institution, which includes so many different things, from languages to political structures. Are such capacious definitions unwarranted? If so, the definitions of a mammal, or of an animal, or of a tree, would also have to be rejected as overly-capacious. Instead of a rejection of the primary classification, capacious outcomes point to the need for sub-classification and further differentiation.

North is often attributed with a definition of institutions that excludes organizations. In key passages that are often cited in support of this claim, North (1990: 3–5, 1994: 361) did not make it sufficiently clear whether he was *defining* organizations as players or treating organizations as players as an *analytic abstraction*. This has created much confusion, with some of his readers insisting that organizations should be *defined* as players. But in personal correspondence – reproduced in Hodgson (2006) – North indicated that he

⁶ The ambiguous terminology of ‘formal’ *versus* ‘informal’ institutions is often deployed but much less often defined. Sometimes ‘formal’ is intended to mean *codified*. Others use it to mean *designed*. Still others use it to refer to *laws*. Given this confusion, the intended meaning of these terms should always be clarified.

treated organizations as players for the purpose of analysis of the socio-economic system as a whole, and that he did not regard organizations as essentially the same thing as players in all circumstances. In saying that ‘organizations are players’, North was making an *abstraction* for particular analytic purposes, rather than *defining* organizations in this way.⁷

When North (1994: 361) assumed that organizations ‘are made up of groups of individuals bound together by some common purpose’, he was less interested in the internal mechanisms by which organizations coerce or persuade members to act together. Crucially, these mechanisms always involve systems of embedded rules. Organizations involve structures, and these cannot function without rules of communication, membership or sovereignty. The unavoidable existence of rules within organizations means that, by North’s own definition, organizations must be regarded as a type of *institution*. Indeed, in correspondence North accepted that organizations themselves have internal players and systems of rules, and hence by implication that organizations are a special type of institution.⁸

As North acknowledged, it is possible for organizations to be treated as actors in some circumstances *and* generally to be regarded as institutions (Hodgson, 2006). Individual agents act within the organizational rule-system. In turn, under some conditions, organizations may be treated as actors within other, encompassing institutional rule-systems. There are multiple levels, in which organizations provide institutional rules for individuals. Under some conditions these organizations can also be treated as actors within broader institutional frameworks (Coleman, 1982, Hindess, 1989). For example, the individual acts within the nation, but in turn the nation can sometimes be treated as a singular actor within an international framework of rules and institutions.

Given that institutions are systems of rules, we need to clarify the nature of a rule. As a first and very simple approximation, we can understand a rule as a learned and mutually understood injunction or disposition, that in circumstances

7 While North in personal correspondence (Hodgson, 2006) agreed that organizations may be treated as a special type of institution, North *et al.* (2009, 15) subsequently contradicted this: ‘In contrast to institutions, *organizations* consist of specific groups of individuals pursuing a mix of common and individual goals through partially coordinated behaviour’. If they had put ‘In contrast to *other* institutions, organizations ...’ then the contradiction would have disappeared. But it is misleading to define organizations as ‘specific groups of individuals’ because specific individuals may leave or enter the organization without the organization changing its identity. North *et al.* (2009: 16) went on to say that ‘most organizations have their own internal institutional structure,’ which confusingly suggested that most organizations contain institutions.

8 Consider Ménard’s (1995) attempt to separate organizations from institutions. He defined institutions as ‘*stable, abstract and impersonal rules, crystallized in traditions, customs or laws*’ (p. 167). Ménard (like others, including myself) also regarded markets as institutions. He continued: ‘Organizations, and especially firms, are specific institutional arrangements, different from those of markets’ (p. 172). This created a tangled terminology where an organization is an ‘institutional arrangement’ but not an institution. He also wrote: ‘I see both markets and organizations as firmly *embedded* in institutions’ (p. 174). But if markets are embedded in institutions, and also themselves institutions, then why cannot organizations be institutions too?

X do Y. In turn, ‘do Y’ must be interpreted broadly, to include prohibitions as well as obligations (Crawford and Ostrom, 1995). The ‘if X do Y’ formulation applies to constitutive as well as procedural rules: hence the ‘do’ can apply to understandings or assignments of status functions, as well as to other actions (Searle, 1995, 2005).

The rules that make up institutions must be more than mere declarations by some authority. As Ostrom (1990, 2000) insisted, they must be rules in actual or potential use in a community, and not merely rules in form. Even if the rule is never violated it must act as a real constraint. There must also be some commitment in the community to follow the rule (Gilbert, 1989). Rules include norms of behaviour and social conventions, as well as legal or formal rules. Searle (1995, 2005) argued forcefully that institutional rules have a constitutive and deontic character, and that these particular kinds of rule are important for institutions.

This raises the question whether we can adequately understand the origin, operation and persistence of institutional rules with agents conceived as utility-maximizing individuals. Doubt has been cast on this elsewhere, stressing that agents sometimes follow rules because of duty or morality, rather than mere utility (Hodgson, 2013; McCloskey, 2015; Merrill and Smith, 2007; Searle, 2001; Sen, 1982). Such concerns bring us back to Smith (1759), who did not adopt an utility-maximizing view of agency (Sen, 2010; Smith, 2013). There is also the question of the guiding role of habit in rule-following (Dewey, 1922; James, 1890; Veblen, 1914). These concerns would not necessarily affect the definition of institutions, but they would lead to vital questions and investigations concerning their nature and operation.

References

- Aoki, M. (2001), *Toward a Comparative Institutional Analysis*, Cambridge, MA: MIT Press.
- Bhaskar, R. (1975), *A Realist Theory of Science*, Leeds: Leeds Books.
- Binmore, K. (1994), *Playing Fair: Game Theory and the Social Contract. Volume 1*, Cambridge, MA: MIT Press.
- Binmore, K. (1998), *Just Playing: Game Theory and the Social Contract. Volume 2*, Cambridge, MA: MIT Press.
- Coleman, J. S. (1982), *The Asymmetric Society*, Syracuse: Syracuse University Press.
- Crawford, S. E. S. and E. Ostrom (1995), ‘A Grammar of Institutions’, *American Political Science Review*, 89(3): 582–600.
- de Queiroz, K. (2005), ‘Ernst Mayr and the Modern Concept of Species’, *Proceedings of the National Academy of Sciences*, 102(supplement): 6600–6607.
- Dewey, J. (1922), *Human Nature and Conduct: An Introduction to Social Psychology*, New York: Holt.
- Gilbert, M. (1989), *On Social Facts*, London and New York: Routledge.
- Harré, R. and E. H. Madden (1975), *Causal Powers: A Theory of Natural Necessity*, Oxford: Basil Blackwell.

- Hindess, B. (1989), *Political Choice and Social Structure: An Analysis of Actors, Interests and Rationality*, Aldershot: Edward Elgar.
- Hindriks, F. and F. Guala (2014), 'Institutions, Rules, and Equilibria: A Unified Theory', *Journal of Institutional Economics*, published online: 16 October 2014, 1–22. doi: [10.1017/S1744137414000496](https://doi.org/10.1017/S1744137414000496).
- Hodgson, G. M. (2006), 'What Are Institutions?' *Journal of Economic Issues*, 40(1): 1–25.
- Hodgson, G. M. (2013), *From Pleasure Machines to Moral Communities: An Evolutionary Economics without Homo Economicus*, Chicago: University of Chicago Press.
- Hodgson, G. M. (2015), *Conceptualizing Capitalism: Institutions, Evolution, Future*. Chicago: University of Chicago Press in press.
- James, W. (1890), *The Principles of Psychology*, Volume 2, New York and London: Holt and Macmillan.
- Kemp, T. S. (2005), *The Origin and Evolution of Mammals*, Oxford and New York: Oxford University Press.
- Knight, J. (1992), *Institutions and Social Conflict*, Cambridge: Cambridge University Press.
- Lawson, T. (1997), *Economics and Reality*, London and New York: Routledge.
- Mantzavinos, C. (2001), *Individuals, Institutions and Markets*, Cambridge and New York: Cambridge University Press.
- McCloskey, D. N. (2015), 'Max U vs. Humanomics: A Critique of Neo-Institutionalism', *Journal of Institutional Economics*, forthcoming.
- Ménard, C. (1995), 'Markets as Institutions *versus* Organizations as Markets? Disentangling Some Fundamental Concepts', *Journal of Economic Behavior and Organization*, 28(2): 161–182.
- Merrill, T. W. and H. E. Smith (2007), 'The Morality of Property', *William and Mary Law Review*, 48(5): 1849–1895.
- North, D. C. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge and New York: Cambridge University Press.
- North, D. C. (1994), 'Economic Performance through Time', *American Economic Review*, 84(3): 359–367.
- North, D. C., J. J. Wallis and B. R. Weingast (2009), *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*, Cambridge: Cambridge University Press.
- Nussbaum, M. C. (1992), 'Human Functioning and Social Justice: In Defense of Aristotelian Essentialism', *Political Theory*, 20(2): 202–246.
- O'Neill, J. (1998), *The Market: Ethics, Knowledge and Politics*, London and New York: Routledge.
- Ostrom, E. (1990), *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge: Cambridge University Press.
- Ostrom, E. (2000), 'Collective Action and the Evolution of Social Norms', *Journal of Economic Perspectives*, 14(3): 137–158.
- Popper, K. R. (1990), *A World of Propensities*, Bristol: Thoemmes.
- Porpora, D. V. (1987), *The Concept of Social Structure*, New York and London: Greenwood Press.
- Robinson, R. (1950), *Definition*, Oxford: Clarendon Press.
- Rowe, T. (1988), 'Definition, Diagnosis, and Origin of Mammalia', *Journal of Vertebrate Paleontology*, 8(3): 241–264.
- Schotter, A. R. (1981), *The Economic Theory of Social Institutions*, Cambridge: Cambridge University Press.

- Searle, J. R. (1995), *The Construction of Social Reality*, London: Allen Lane.
- Searle, J. R. (2001) *Rationality in Action*, Cambridge, MA: MIT Press.
- Searle, J. R. (2005), 'What is an Institution?' *Journal of Institutional Economics*, 1(1): 1–22.
- Sen, A. (1982), 'Rights and Agency', *Philosophy and Public Affairs*, 11(1): 3–39.
- Sen, A. K. (2010), 'Adam Smith and the Contemporary World', *Erasmus Journal for Philosophy and Economics*, 3(1): 50–67.
- Smith, A. (1759), *The Theory of Moral Sentiments; or, An Essay Towards an Analysis of the Principles by which Men Naturally Judge Concerning the Conduct and Character, First of their Neighbours, and Afterwards of Themselves*, London and Edinburgh: Millar, and Kincaid and Bell.
- Smith, V. L. (2013), 'Adam Smith: From Propriety and Sentiments to Property and Wealth', *Forum for Social Economics*, 42(4): 283–297.
- Sober, E. (1980), 'Evolution, Population Thinking and Essentialism', *Philosophy of Science*, 47(3): 350–383.
- Veblen, T. B. (1914), *The Instinct of Workmanship, and the State of the Industrial Arts*, New York: Macmillan.