JESÚS ZAMORA BONILLA

SCIENCE AS A PERSUASION GAME. AN INFERENTIALIST APPROACH*

ABSTRACT

Scientific research is reconstructed as a language game along the lines of Robert Brandom's inferentialism. Researchers are assumed to aim at persuading their colleagues of the validity of some claims. The assertions each scientist is allowed or committed to make depend on her previous claims and on the inferential norms of her research community. A classification of the most relevant types of inferential rules governing such a game is offered, and some ways in which this inferentialist approach can be used for assessing scientific knowledge and practices are explored. Some similarities and differences with a game-theoretic analysis are discussed.

1. Introduction

Unlike Trappist monks, who live under a vow of silence, scientists experience a furious frenzy of verbal communication. The volume of scientific 'literature', including books, journals, conference proceedings, pre-prints, abstracts, and laboratory reports, has grown exponentially since the times of Copernicus. This accumulation may seem anti-economical, especially because much of what is published will never be read by anybody (besides the referees and journal editors) or will be quickly dispatched with just a superficial glance.[1] But I will not try to analyse here whether scientific writing is economically efficient or not. (An argument will be offered, nevertheless, that addresses the role of 'unread' papers in scientific communication.) My main goal is to outline an approach towards understanding scientific research that takes into account that *communication* is one of its essential elements. On this approach, scientists are seen as engaged in a type of *game* whose main purpose consists in the *persuasion* of their colleagues. Some attempts have been made to analyse the behaviour of researchers from a game-theoretic point of view (e.g., Blais (1987), Bicchieri (1988), Goldman and Shaked (1991), Kitcher (1993), Luetge (2004), Zamora Bonilla (forthcoming a)).[2] But here I will explore the idea that scientific research is a *language game* as originally defined in the philosophy of language and pragmatics since Wittgenstein (1953). My preferred approach to the theory of language games is Robert Brandom's *inferentialism* (Brandom 1994, 2000), which integrates many developments of previous theories. The structure of this paper is as follows. In section 2, I describe the basic elements of an inferentialist theory of language, the central one being the notion of an 'inferential rule', and offer a brief comparison with game theory. Section 3 sketches an inferentialist reconstruction of research as a persuasion game. Section 4 presents a classification of the most general types of inferential rules that constitute the game of science. Section 5 discusses the origin and evolution of scientific norms. Section 6 explains how an inferentialist approach can be put to the service of a normative assessment of scientific knowledge.

2. Language games and human action: an inferentialist approach

The notion of a 'language game' has attracted the attention of scholars from a wide variety of disciplines, but there is hardly a systematic and unified account of this notion. This is in contrast to mathematical game theory, where a wide agreement exists about the basic elements of a 'game'.[3] My choice of Brandom's inferentialism as a point of departure therefore seems to require at least some justification, though a full defence of it would demand more than a single paper. Brandom describes his theory as an 'inferential semantics grounded on a normative pragmatics'. The 'inferential semantics' part means this: Instead of understanding the *meanings* of linguistic expressions as based on *representations* of aspects of the world by pieces of speech or text, those meanings are primarily explained through the *inferential links* that each expression has with others. The 'normative pragmatics' part refers to the idea that human *actions* (not only verbal ones) are understood as structured and motivated by the *deontic statuses* of the agents. The semantic part being 'grounded' on the pragmatic one means that it is not the semantic properties of language that determine what people can do with it, but rather the other way around. The most important aspects of Brandom's theory employed here are the following:

a) To each participant in a conversation we can attribute a *deontic score*, i.e., an indication of the things that, at every moment, she is 'committed' or 'entitled' to do (or not to do). Many, though by no means all, of these entitlements and commitments refer to what one can, or has to, *assert*.

b) The agents' normative scores evolve according to the *inferential rules* that are accepted within the community of speakers. These rules specify how each score is modified by the actions and assertions of each agent, as well as by other events.

c) Interaction between the agents basically takes the form of 'scorekeeping',[4] i.e., not only by every speaker *undertaking* specific commitments or entitlements, but also by her *attributing* them to others.

We can attempt to summarise these three elements in a single formula. Consider a given speech community of n members. Let D be the set of all possible deontic scores this community's language can express, and assign a number to each element of D. Let d be an n-dimensional vector, assigning to each individual one of the possible scores. Let E be a collection of possible *events*, including the possible *actions*, A, each individual can perform (including assertions); E can also contain the 'null' event, i.e., one asserting that nothing new happens. Then we can simply represent the inferential rules of a speech community as a function of the form $D^n \times E \longrightarrow D^n$. Inferential rules are thus the 'kinematic laws' of deontic scores.[5] Lastly, we would need something like a 'dynamic law', expressing the probability that a particular action is performed by an individual, given her and probably others' deontic scores. (Formally, this 'law' would be a probability distribution over A conditioned by the vectors in D^n.)

This general framework suggests modelling social phenomena as a type of Markovian process. The main difference from a conventional game-theoretic analysis is that, in game-theoretic models, rational strategic choices are the 'force' behind the decisions of individuals, whereas inferentialist models assume that people tend to do what they 'have' to do. I will comment in the next paragraph on why much of this difference is more apparent than real, but I want to concentrate now on an aspect of social action that inferentialism can illuminate better than game-theoretic or rational-choice models. It is the fact that *rational* action is basically action derived from a process of *reasoning*. Rational-choice models simply 'assume away' the real process of deliberation that agents carry out within their own minds or within public conversation, and replace a description of this *process* by the assumption that its *output* corresponds to the one that maximises a utility function given beforehand. This has proved to be a very powerful heuristic, but it has also encountered serious limitations, especially when applied to situations where normative concerns are paramount.[6] Inferentialist models, in contrast, assume that actions derive basically from *commitments*, and can thereby analyse the way these commitments are 'constructed' through a deliberation process. This also allows a normative analysis of the process itself, i.e., of the *arguments* by which the decision is reached. This is particularly important in the case of science, where argumentation is essential.

The contrast between inferentialist and game-theoretic approaches is not absolute, however. In the first place, individual behaviour is often not fully determined by commitments. There are many reasons for this indeterminacy: agents can be entitled to choose between several options; commitment usually comes in degrees; someone can be committed to perform two incompatible actions; and agents can decide to break their commitments sometimes. In all these cases, there is some room for 'strategic' choice. In principle, it is possible to analyse these choices game-theoretically by introducing the weight of normative commitment as an element of the players' pay-offs.[7]

In the second place, we can simply interpret an agent's *patterns* of reasoning and behaviour (i.e., the inferential norms followed by her) as a strategy in her game 'against' the other members of her community. In this case, it is reasonable to assume that patterns that often lead to non-Nash-equilibrium situations will tend to be substituted by other patterns. This also justifies the applicability of game-theoretic tools, particularly those of evolutionary game theory, where agents don't choose their 'strategy', but are rather assumed to be 'programmed' by it. A different way of expressing the condition that the norms used in a scientific community must regularly lead to Nash-equilibrium social states is to say that these norms must be 'incentive-compatible'.[8]

Moreover, we should have an overtly *instrumental* position towards the game-theoretic and the inferentialist paradigms. The most important point should not be whether they are 'right', 'true to the facts', or not (a difficult question, to say the least), but to what extent they generate illuminating models of social facts. Some recent developments in the philosophy of science point towards the virtue of such a tolerant attitude (e.g., Cartwright 1999, Nickles 2002).

3. Playing the game of science.

Like most language games, scientific communication is systematically oriented towards persuasion, but, unlike other common language games, scientific communication is one in which the players' goal is not primarily to convince others *to do something*, but to make them *accept some propositions*. This does not mean that *practical actions* are irrelevant in the game of science. On the contrary, they are an essential part of it, but inferentialism suggests that we should consider the role of scientists' *material practices* in their essential connection to the inferential rules that govern their *communication practices*. On this view, each scientist's main goal (but in no way her only one) is *recognition*, i.e., *she wants her colleagues to explicitly accept the claims advanced by her* – claims that may have the character of theories, laws, models, facts, theorems, and so on. Of course, most scientists don't pursue 'bare' fame, but aspire to get recognition '*for* having made an important discovery'. I suggest interpreting the phrase in inverted commas as the thesis that it matters to scientists that the inferential rules according to which they gain recognition are 'appropriate' ones. (I shall return to this issue in section 6.) Perhaps some researchers do not care about recognition at all, but I will assume that, as a matter of fact, most do, and, more importantly, that scientific practice is basically organised around the pursuit of this goal. The most important problems for any recognition-seeking researcher, and the reason why game-theoretic considerations are important, are, first, that the scientist's colleagues are all *simultaneously* trying to do the same, and second, that the choice of accepting or rejecting one's presumed discoveries is always made *by one's colleagues*.

To the rescue of our unfortunate researcher comes the fact that communication takes place according to a system of inferential rules that, amongst other things, determine what claims somebody is *committed* to make given what other propositions have been accepted by him before, or whether he has the obligation of accepting them 'by default' (e.g., any chemist is obliged to accept the claims contained in the standard periodic table). So, what a researcher may try to do is to find some claims that, according to these inferential rules, and taking into account the commitments she and her colleagues have undertaken before, *a)* she is *entitled* to make, and *b)* her colleagues are *committed* to accept. *Scientific communication can be understood, then, as the set of moves each researcher makes in this game of looking for 'successful' claims*.[9] In the next sections, I will describe how the inferential rules of a scientific community can make epistemically interesting a game like this. In the remainder of this section, I will concentrate on a different problem, directly related to the question with which I began, i.e., 'why do researchers write so much?'.

In general, scientific norms do not allow anyone to accept two contradictory claims about the same issue (or more than one 'solution' to the same problem). This entails that, for every claim accepted by some researcher, there will be many other claims that she does not accept, and the competitive process of science entails that most of the claims that have been proposed will be nearly universally rejected or ignored. If competition is very strong, and if the norms are so demanding that acceptable solutions to interesting problems are very difficult to discover, it may easily happen that many researchers are statistically condemned to never making any important 'discovery' (see note 1). Hence, if scientific research has some appeal for recognition-seeking individuals, at least some other recognition-related goal attainable by them must exist, different from the public acceptance of the *validity* of their discoveries. My suggestion is that this secondary form of recognition consists in the public acceptance of some pieces of work, *not as 'right', but as 'good'*, i.e., not as a *claim* it is compulsory to accept, but as an *argument* that is coherent with the research community's inferential practices.

In real science, this secondary form of scientific merit usually takes the form of having your papers *published* in good journals, whereas the primary form of recognition consists in the content of those papers being *cited and used* by your colleagues in further papers. Usually, the better your papers are, the more probable it is that the claims you have made in them become accepted, but this is certainly not a necessary implication. First, a different solution to the problem you have tackled, still better than the one proposed by you, may be invented. (This will not necessarily reduce your work's scientific quality, but will make fewer people accept it.) Second, you may happen to defend, based on reasons that are not particularly good, a very good idea, which later becomes widely accepted. (Perhaps you are so lucky that better arguments in defence of your hypothesis are provided later by other people.) I have introduced the term *external score* to refer to the first, fundamental type of recognition, and *internal score* to refer to the merit derived from the coherence of a scientist's inferential practices with the rules of her community (Zamora Bonilla, forthcoming a). A scientist's *global score* is some aggregation of both. The particular form of aggregation depends on her own subjective valuation of both types of recognition, though it is reasonable to assume that the weight attached to the external score is much higher than that of the internal score.[10] The game of science proceeds, then, by each researcher permanently trying to maximise her global score (which will often lead her to favouring strategies enhancing the internal score), and this will have as an unintended consequence the determination of her colleagues' external scores, as the researcher happens to be committed to accept or to reject the claims proposed by them.

As we have just seen, the fact that a researcher's global score is constituted by two different elements entails that, in some cases, there will be more than one possible strategy to increase it. For example, one can either try to elaborate bold hypotheses, or decide to work within a more conservative project. This possibility suggests an interesting case for the application of game theory, as the reward a scientist can expect to gain from her decision will depend on the choices of her colleagues: if many of them choose an *internal* score optimisation strategy, i.e., if they are tightly bound to accept any conclusion that you *prove* from their previous assumptions, attaining a high external score will be easier for you, and hence it can be more profitable to choose an *external* score optimisation strategy. Figure 1 represents how a social equilibrium is determined when this kind of interdependence occurs. The scale of the graph represents the average ratio between the effort devoted by each researcher to external versus internal score enhancing activities. The function *f* maps each possible value of that ratio into the average ratio that would emerge after every researcher decides what strategy is best for her given the average behaviour. The equilibrium r^* is the 'fixed point', i.e., the point at which the reaction function crosses the identity line, i.e., when the community's aggregate response to a given social state is just this state.[11] Only at this point are all the chosen strategies mutually compatible (which does not mean that all researchers will be choosing the same strategy). Of course, the actual equilibrium cannot be determined *a priori*, as it will depend on the particularities of each scientific field and context. It would be interesting to investigate empirically whether the *f* function differs between different research areas, countries, or times, and, if it does, what factors these differences depend on.

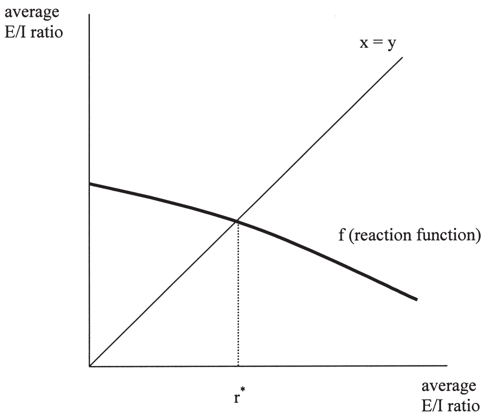

FIGURE 1

4. A general classification of the inferential rules of science

Seeing the inferential norms of any language game as 'kinematic laws' of the deontic scores suggests a straightforward way of classifying these norms:

(1) Some norms (*argumentation rules*) simply change an existing score into a different one (e.g., they may add or remove a commitment), independently of whether something else happens. This can be represented by placing the null event in the left part of the formal expression of these rules,[12] and the obligation or permission to make some specific claim as the output of the rules. Logical norms are a prime example of argumentation rules, but context-dependent patterns of material inference also belong into this category.

(2) Other norms (*entry rules*) determine how the normative scores are affected by some events different from the mere existence of previous commitments and entitlements. An example is given by the rules which determine what claims one is entitled or committed to make depending on one's own perceptual experiences and those of others. These norms express how 'the world' affects the things we are entitled or committed to assert or to deny. It is important to take into account that, in general, non-linguistic events will only generate some new deontic status on the basis of some *previous* state of the deontic scores.

(3) Lastly, some norms (*exit rules*) determine how the claims included in the normative score of one speaker (i.e., the propositions she accepts) transform into the obligation to perform, or to abstain from performing, some non-verbal actions. These are the rules of practical reasoning, i.e. those determining how deliberation leads to decision making.

In the case of science, some further distinctions within each type of norm are useful. First, among *argumentation rules*, we can distinguish between norms for the *evaluation* of claims and norms for the *choice* of claims. The former express the possible virtues and flaws of scientific claims (theories, hypotheses, empirical reports, and so on); these norms allow us to determine the 'scale' against which the quality of a scientific claim can be *assessed*. The latter are those rules that command a researcher *to accept or to reject* specific statements according to their quality; these norms can also be interpreted as determining 'how well justified' the acceptance or rejection of a particular claim is. In a nutshell, norms for the choice of claims determine *how good* the accepted intellectual products of scientific research are, whereas norms for the evaluation of claims determine *in what sense are they 'good'* (see Zamora Bonilla 2002).

Second, regarding *entry norms* we can distinguish rules for *evidence gathering* from rules about *authority*. The former determine what claims a researcher is entitled or committed to do after the occurrence of some 'publicly observable event' (it is important to notice that the inferential rules also determine what can legitimately count as such an event). The latter express what claims a scientist is entitled or committed to adopt because some *other* scientists have adopted them before. Hence, rules for evidence gathering have the role of determining the *appropriate empirical methods*, and rules about authority serve to identify who are the *experts* about what. Both types of rules are essential for an efficient *economy of trust* within scientific communities.[13]

Lastly, exit norms can be interpreted as rules for *resource allocation*, and so they can be classified according to the different types of resources employed in scientific research: space for publication in journals, distribution of research effort within teams, distribution of money among research projects, grants, prizes, jobs, and so on. It is important to notice that many actions (e.g., experiments) produce new events, which, in turn, change the deontic scores.

Figure 2 summarises how the rules of a scientific community work. Starting from a given ('exogenous') state of the researchers' deontic scores (D_x), the occurrence of some events (E) triggers its transformation into a new ('endogenous') state (D_n). The latter commands the performance of some actions (A), which generate new events. These, in turn, change again the endogenous state, which can also be transformed by the application of transformation rules, and so on. The presence of D_x reminds us that no research starts in a normative vacuum, and thus, new events can only attain some meaning (i.e., normative inferential strength) through the research community's previous commitments. (But this does not entail that the old commitments will necessarily be preserved by the operation of the research process.) In the end, all this will go on until a *corpus* of 'knowledge' is produced, which I shall identify here just with the commitments that become generalised in the deontic scores of the members of the research community, particularly when an 'equilibrium' state is reached, i.e., when the inferential rules do not lead us to expect new actions or events to change the deontic scores. Note that, according to inferentialism, it is not necessary that an agent explicitly asserts a claim for being really committed to it; it is not even necessary that she believes in its truth. For her to have a particular commitment, it is enough that it follows according to the inferential rules accepted in the relevant community and from her other commitments. So, it is not necessary that every member of the scientific community overtly *accepts* a claim for this claim to constitute 'knowledge' in the sense expressed above.[14]

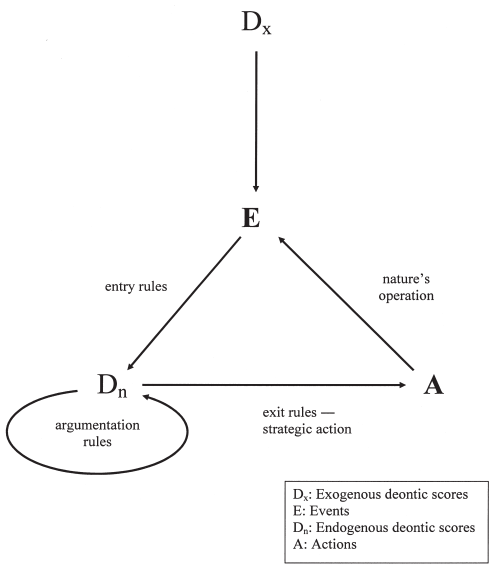

FIGURE 2

5. The origin and evolution of scientific norms

The next important question is why scientific rules are the ones they are. I may admit that some inferential norms can be universal, but I am pretty sure that most of them are contingent and context dependent. Different communities may have different practices, and these practices can change as well, because most rules depend on a community's knowledge, beliefs, instruments, or even institutions, and all these factors change permanently. In the next section I will show why the recognition of this variability is not an invitation to relativism nor to methodological anarchism. Now I will discuss the mechanisms determining that certain specific rules are actually in force in a scientific community. Curiously, although Robert Brandom has attributed a central role to the *community* as the holder of a language's inferential rules, he has not discussed why diverse linguistic communities have such different inferential rules, i.e., why they have different *concepts*. Brandom's strategy seems to be, instead, to scrutinise which inferential rules constitute a 'possibility condition' for any conceivable language game. Something similar could also be done in the case of science, and later in this section I will make some suggestions in this direction. But it is interesting to reflect first about what *general* factors determine the adoption of some *specific* rules rather than others.

From an individual scientist's point of view, it is true that inferential norms are given (as in most language games). Actually, by constituting the *meanings* of her claims and actions, the existing norms will determine even what a single researcher is able to think and understand, and so it is unlikely that she even considers the possibility of changing those norms. But it is also a fact that rules *do* change. How is this possible? The recognition of this apparent paradox is what led Thomas

Kuhn (1962) to suggest that paradigm shifts were more like sudden 'religious' conversions than like the outcomes of rational arguments. A more rationalistic account can be offered, however, within an inferentialist perspective. The following facts are relevant for that account:

(a) Not all inferential rules are equally *strict* in determining understanding and action. Some rules are taken by the agents themselves as conventional, and hence more easily subject to variation. Even in mathematics, we can draw a continuum of 'inferential strength' from rules like those contained in a multiplication table to those relating to 'appropriate' heuristic methods, or those of statistical inference. From Brandom's inferentialism, the starting point of the analysis is not the 'inherent logical force' of the rules, but the way in which the community negotiates the *normative significance* of behaviour more or less incoherent with the accepted norms. We can then restate this first point by saying that, within each scientific community, not every violation of the norms will be considered equally *reprehensible*.

(b) The existing norms do not necessarily constitute an *exhaustive*, nor even a *consistent* system. In many situations the norms can fail to determine the next state of the deontic scores, or can lead to contradictory commitments. When this happens, it is reasonable that agents react by considering the possibility of revising some rules. Again in Kuhnian terms, a way to solve a 'paradigm crisis' is to eliminate those concepts whose use led to 'anomalies'.

(c) Actual inferential norms are usually not like algorithmic programmes (although they *seem* like ones when represented in a mathematical model). Each individual agent has to construct, from her own experience, a subjective, usually tacit, *interpretation* of what the 'right' rules are, and it may happen that the subjective interpretations of different individuals are not entirely coherent. An agent can also realise that she was not understanding 'properly' some rule. In cases like these, negotiations to determine the 'right' interpretation of the norms will arise. Again, this is illustrated by cases of Kuhnian 'scientific revolutions', though this does not entail that the resolution of the crisis can not be the result of *reasoned* arguments. (In some cases, the working of conversation in figure 2 can lead to the adoption of a new rule or the rejection of an old one.)

(d) As in many other social contexts, in science the main goal of the training process is to endow young scientists with an 'appropriate' mastery of the inferential norms. However, this process of 'replication' of the rules from teachers to students, or from researchers to colleagues, is not necessarily perfect and can give rise to 'mutations'.

(e) Lastly, some norms can be more *beneficial* than others for the members of the community, or for some of them. This is obviously true for exit rules, i.e., for those that determine the allocation of resources and the distribution of benefits, but it may also happen with other types of norms. For example, it makes an enormous difference for recognition-seeking researchers what rules for theory choice are accepted within their community, since each researcher's probability of discovering an 'acceptable' theory will depend on what specific rule is actually established (cf. Zamora Bonilla, 2002).

These properties of the inferential norms suggest that there can be two general mechanisms for their origination or transformation: *conventional agreement* and *evolution*. Scientists can simply meet (physically or not) and decide that some things must be interpreted or done in such and such a way, or, more usually, they can delegate such a decision to a small number of experts. Consensus congresses, constitutions of scientific societies, and international standards or measurement agencies, are examples of this type of conventional setting of the inferential norms. A useful approach to analyse the collective choice of scientific norms is Constitutional Political Economy (see, e.g., Brennan and Buchanan 1985), which can be described as the application of game theory to decisions about what game to play, i.e., what 'constitutional' norms to adopt.[15] But evolutionary processes, in which every scientist tries to adapt her inferential practices to those she observes in her most successful colleagues, are probably more frequent than explicit agreements. In this case, several theoretical approaches are available for modelling the evolution of norms (see Vromen 1995 and Nelson 2002 for a recent survey; Sperber 1996 and Blackmore 2000 offer two interesting evolutionary approaches to the study of cultural items). Nevertheless, as individual decisions under both mechanisms are driven by how beneficial the adoption of a particular rule is for each agent, given the rules adopted by others, it seems that a reasonable minimal requirement is that the chosen norms constitute a Nash equilibrium. Again, this does not entail that all agents will choose the same norms in equilibrium, only that nobody can make a better choice given the choices of the rest.[16] Taking this into account, the most important step in an analysis of scientific rules is to identify how each possible norm affects the 'payoffs' of the scientists whose behaviour will be regulated by it, but this connects directly with the question about the evaluation of scientific norms, which is the topic of the next section.

6. Assessing the game of persuasion

Our reconstruction of science as a persuasion game opens several avenues for the normative evaluation of scientific knowledge and scientific practice, either in specific case studies, or regarding general features of science. First, there are two broad categories of *items* to be evaluated: the global processes of interaction between scientists (or between them and other agents), and the particular pieces of argumentation they are using. In other words, we can make a normative assessment of the *game* a group of scientists are playing, or we can assess the particular *moves* they are making within the game. Second, we can adopt the *stance* of an internal participant in the game, or of an external observer. Figure 3 summarises the main general questions that can be raised under each one of the four resulting possibilities. Regarding the 'Arguments' column, by adopting a participant's perspective we might ask whether the behaviour of the scientists has obeyed in an appropriate way the inferential rules they *themselves* have adopted. We would adopt in this case a position similar to that of the referees in a sport game (hypothetically, at least). On the other hand, by adopting an external observer's point of view, we could judge whether those rules seem rational *to us* or not; for example, would we have preferred that the game had been played according to different rules? Regarding the 'Games' column, the main difference between the two different points of view consists in whether we take into account only the interests of the participants in the game (i.e., those people having some choice to make in it), or the interests of other agents as well. In the first case, once we reconstruct a scientific episode as a game, an obvious question is whether the choices of the participants constitute a Nash equilibrium or not; in the latter case, we would be committed to assert that at least one of the participants has behaved in an irrational way. Another important question is whether the situation is *efficient* or not, i.e., whether the agents would have got a better pay-off by (more than one of them) having made some different choices.[17] Lastly, regarding those agents that are 'alienated' from the game, in the sense of having no significant power of choice, we might then investigate under what rules the outcome of the game would have been more efficient for them.

What is evaluated? *Point of view*	*Arguments*	*Games*
Internal participant	Are the rules applied in an appropriate way?	Is the situation an equilibrium? Is it efficient?
External observer	Do *you* find the rules and the arguments persuasive?	How does the situation affect other people?

FIGURE 3

Of course, formulating relevant normative questions of this kind and putting forward appropriate answers is by no means an easy task, and nothing in the former classification saves the work of gathering an exhaustive knowledge of the particularities of those cases we might decide to assess. The main point in having a generic schema like the one I have just depicted, consists in showing that such an evaluation is a *legitimate* enterprise, for it follows in an absolutely natural way from the understanding of the scientific process as a game played according to some rules. A metaphor can be useful here. According to the inferentialist approach, the institutions most similar to scientific research are probably *sports*: both in sports and in science, individuals compete, either alone or in teams, to attain public recognition by performing in the best possible way under the constraints of some rules that, by providing the 'kinematic laws of scorekeeping', define what success consists in. This competitive and regulated nature of sports entails that the most important thing we can do with them (besides playing them, of course) is anything but *judging* them.

In this sense, the approach offered in this paper fits nicely with other approaches like those of Goldman (1999) and Koppl (2006), in which the primary goal of epistemological thinking is to assess the *efficiency* of cognitive practices, considered as social mechanisms for the production of knowledge. The main difference to those approaches would be that inferentialism allows for a more liberal and flexible range of evaluative stances. First, it does not necessarily assume that there is some unique epistemic goal to be attained by scientific practices, nor that epistemic goals are the only essential ones. Second, conflict and complementarity is recognised not only within the goals of a single agent, but also between the goals of different agents, as it is evident in those cases demanding a game-theoretic reconstruction. And, third, a distinction is clearly established between, on the one hand, the assessment of a scientific practice with the standards of its participants and, on the other hand, an evaluation based on *our own* preferred goals.

I shall end by discussing why, in spite of its liberalism, the inferentialist approach is not committed to a relativistic interpretation of scientific knowledge. The following are some of the reasons why inferentialism can be seen as an *objectivist* approach (they are presented according to the four normative stances in figure 3):

(a) *(Internal criticism of specific arguments)*. In contrast to Feyerabend's famous slogan (Feyerabend 1975), in the game of science not anything goes. Inferential rules have real normative cogency for players, and so, when we evaluate whether a scientist has made an 'appropriate' move according to *her own* commitments, our answers will be objectively right or wrong. This is not essentially different from when we judge whether a particular action in a football match has been a foul or not, though perhaps the same action would have been correct in other sports (say, boxing), or in the same sport once the rules change (say, offside). It is true that successful research sometimes violates some methodological norms, but this can only be so if other collectively accepted norms specify that this behaviour constitutes a success.[18]

(b) *(Internal criticism of the game)*. Regarding inefficient situations, their being so is also an objective question: given the goals and preferences of the agents engaged in a particular social situation, it may simply happen that a different 'solution' would have been preferred by all the players. Scientific assessment in this case amounts to engaging in the same kind of arguments and negotiations that the participants themselves would address if they wanted to *successfully* move towards a more efficient situation.

(c) *(External criticism of the game)*. Again, this kind of evaluation consists in objectively describing what specific benefits or disadvantages people different from scientific players themselves experience as a result of the research process, according to the actual values of the former.

(d) *(External criticism of arguments)*. I have left this case for the end, because all the other types of criticisms depended on the (conditional) acceptance of the validity of some inferential norms, according to which the persuasion game was being played. It may seem that a relativist stance is compatible with the three previous possibilities, for, after all, does relativism not amount to the claim that everybody is free to have her own evaluation standards? The kinds of evaluation I have considered until now can apparently be reduced to assertions like 'this is good/right from the point of view of X, and bad/wrong from the point of view of Y', or so it seems. But, actually, most of my examples, particularly in *a* and *b*, were of a subtly different kind, like 'X is acting in a way contrary to the norms *she* is committed to, or in a way contrary to the fulfilment of *her own* interests'. Furthermore, we can also evaluate the specific norms according to which the game of science is played each time. Of course, this evaluation has to be based on the acceptance of some norms of a 'higher level', so to say, and these can be different from the ones the players actually accept. This might seem a clear example of Kuhnian *incommensurability*, but I will try to show that it is not.

In the first place, a claim like 'according to the goals and beliefs of group A, the set of inferential norms N is efficient, but according to the goals and beliefs of a different group B, other norms, M, are more efficient' can be absolutely objective, even if N and M are mutually contradictory. This is simply like comparing the rules of football and those of gymnastics: *both* can be rationally defended, and, what is more important, it is perfectly possible to conceive a rational debate about which of the two sports to play in a specific meeting (assuming there is only time, people, or space for one).

In the second place, since argumentation about rules necessarily makes use of 'higher level' norms, it is likely that the number and extent of the possible criteria according to which this argumentation takes place is severely limited, i.e., it is likely that there are only a few possible 'general scientific values' that serve as a final guide in the construction of scientific arguments, as Kuhn (1977) himself recognised. The conflict between different schools would again reduce, hence, either to their attaching a different weight to each of these values, or to the fact that the specific circumstances of each research process make it easier to satisfy some values than others.

In the third place, it is very likely that longstanding competition between researchers during the last centuries has led science, as a global institution, to develop a set of *general* criteria for the assessment and choice of scientific claims that are extremely efficient in the production of theories, models and data of the highest quality according to *all* the values referred to above, so that *it can be practically impossible to conceive some general criteria that are still better*. Those norms that are more specific and context dependent will surely be more sensitive to the appearance of anomalies within concrete research fields.[19]

But I think that, over and above the recognition that different norms may have different degrees of efficiency in different circumstances (and that this is an objective question), inferentialism helps to defeat relativistic approaches by insisting that *scientific claims are basically commitments supported by reasons*. After all, evaluating a scientific argument from an external perspective amounts to putting the following question: what would have been a more correct way of doing it, according to *you*? So, it is impossible to adopt an inferentialist stance without being committed to the thesis that there are some ways of doing research that are more appropriate than others. Except by helping to identify some common argumentation criteria, inferentialism does not attempt to find out what these 'right' ways of doing science are, but it helps to justify the objectivity of science by making us recognise that the epistemic claims to which we are finally committed are not necessarily the ones we would have wanted to defend in the first place, but the ones our reasoned dialogues have led us to accept in the end.

References

Bicchieri, Christina. 1988. "Methodological Rules as Conventions," *Philosophy of the Social Sciences*, 18: 477-95.

Blackmore, Susan. 2000. *The Meme Machine*. Oxford: Oxford University Press.

Blais, M. J. 1987. "Epistemic Tit-for-Tat," *Journal of Philosophy*, 84: 363-75.

Brandom, Robert. 1994. *Making It Explicit. Reasoning, Representing, and Discursive Commitment.* Cambridge, MA: Harvard University Press.

Brandom, Robert. 2000. *Articulating Reasons. An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.

Brennan, Geoffrey, and James Buchanan. 1985. *The Reason of Rules. Constitutional Political Economy*, Cambridge: Cambridge University Press.

Cartwright, Nancy. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Cole, Stephen. 1992. *Making Science: Between Nature and Society.* Cambridge, MA: Harvard University Press.

Feyerabend, Paul. 1975. *Against Method.* London: NLB.

Goldman, Alvin I. 1999. *Knowledge in a Social World*. Oxford: Oxford University Press.

Goldman, Alvin I., and Moshe Shaked. 1991. "An Economic Model of Scientific Activity and Truth Acquisition," *Philosophical Studies*, 63: 31-55.

Jarvie, Ian. 2001. *The Constitution of Science*, Amsterdam and Atlanta: Rodopi.
Kitcher, Philip. 1993. *The Advancement of Science*. Oxford: Oxford University Press.
Koppl, Roger. 2006. Epistemic Systems. *Episteme* 2 (2): 91-106.
Koppl, Roger, and Richard N. Langlois. 2001. "Organizations and Language Games," *Journal of Management and Governance*, 5: 287-305.
Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press.
Kuhn, Thomas S. 1977. "Objectivity, Value Judgments, and Theory Choice," in *The Essential Tension*. Chicago: The University of Chicago Press.
Lewis, David. 1969. *Convention. A Philosophical Study*. Cambridge, MA: Harvard University Press.
Lewis, David. 1979. "Scorekeeping in a Language Game," *Journal of Philosophical Logic*, 8: 339-359.
Luetge, Christoph. 2004. "Economics in Philosophy of Science: A Dismal Contribution?," *Synthese*, 140: 279-305.
Mäki, Uskali. 2004. "Economic Epistemology: Hopes and Horrors," *Episteme* 1 (3): 211-222.
Nelson, Richard. 2002. "Evolutionary Theorizing in Economics," *Journal of Economic Perspectives*, 2: 23-46.
Nickles, Thomas. 2002. "From Copernicus to Ptolemy: Inconsistency and Method," in Joke Meheus (ed.), *Inconsistency in Science*. Dordrecht: Kluwer, 1-33.
Parikh, Prashant. 2001. *The Use of Language*. Stanford: CSLI Publications.
Rabin, Matthew. 1993. Incorporating Fairness into Game Theory and Economics, *The American Economic Review*, 83: 1281-1302.
Rubinstein, Ariel. 2000. *Economics and Language*. Cambridge: Cambridge University Press.
Sen, Amartya. 1997. "Maximization and the Act of Choice," *Econometrica*, 65: 745-79.
Sperber, Daniel. 1996. *Explaining Culture. A Naturalistic Approach*. Oxford: Basil Blackwell.
Sugden, Robert. 2000. "The motivating Power of Expectations," in Julian Nida-Rümelin and Wolfgang Spohn (eds.), *Rationality, Rules, and Structure*. Dordrecth: Kluwer, 103-130.
Vromen, Jack. 1995. *Economic Evolution. An Inquiry into the Foundations of the New Institutional Economics*. London: Routledge.
Wittgenstein, Ludwig. 1953. *Philosophical Investigations*, Oxford: Basil Blackwell.
Zamora Bonilla, Jesús P. 2002. "Scientific Inference and the Pursuit of Fame: A Contractarian Approach," *Philosophy of Science*, 69: 300-323.
Zamora Bonilla, Jesús P. (forthcoming a). "Science Studies and the Theory of Games," *Perspectives on Science*.
Zamora Bonilla, Jesús P. (forthcoming b). "Methodology and the Constitution of Science: A Game Theoretic Approach," in Max Albert, Dieter Schmidtchen and Stefan Voigt (eds.), *Scientific Competition. Conferences on New Political Economy 24*, Tübingen, Mohr Siebeck.

Notes

* Some of the ideas in this paper were presented at a seminar on 'Methodology and the Constitution of Science', in the Alpbach European Forum, August 2004, and in a seminar on 'Scientific Competition' at Saarland University, Saarbrücken, October 2005 (Zamora Bonilla, forthcoming b). I thank the organisers and participants in those events, as well three anonymous referees of this paper, for their detailed comments. Financial support from Spanish government's research projects BFF2002-03353 and HUM2005-01686/FISO, as well as from Urrutia Elejalde Foundation, is acknowledged.

1 In a classical study, sociologist Stephen Cole documented that a 70 percent of the new researchers in physics receive *no* citations within the five years after their Ph.D., and only a 10

percent get five or more citations. Cf. Cole (1992), pp. 218 ff.

[2] See Mäki (2004) for a discussion of other economic approaches to social epistemology.

[3] Parikh (2001) is perhaps the most systematic attempt of grounding the analysis of language games on game theory itself. Rubinstein (2000) also uses game theory to illuminate some aspects of linguistic communication. Koppl and Langlois (2001) is one of the few applications of language games to economic theory.

[4] This idea was advanced by Lewis (1979).

[5] In a slightly more sophisticated way, we might represent deontic scores not as vectors, but as $n \times n$ matrices, in which the element d_{ij} would represent the deontic score that individual *i* attributes to individual *j*. Similarly, the elements of *E* could be substituted by ordered sets of *n* members, with e_i standing for the event *e* as interpreted by agent *i*.

[6] Cf., e.g., Sen (1997).

[7] Cf., e.g., Rabin (1993); but see Sugden (2000), pp. 124 ff., for an argument against the reducibility of normative motivations to classical game theory

[8] I owe this appreciation to an anonymous referee.

[9] There can be 'defensive' moves as well, i.e., one can give up a previous commitment to avoid accepting a particular claim.

[10] Just for simplicity, a scientist's internal score could be operationalised as a function of her *published* papers, whereas her external score would be a function of the favourable citations she receives.

[11] A sufficient condition for the existence of equilibrium is that *f* is continuous and that, for some value of *x*, *f(x)* equals 0.

[12] Recall that inferential rules were functions of the form $D \times E \longrightarrow D$.

[13] This can be related to the distinction introduced in Kitcher (1993, ch. 5, § 7) between 'encounters with nature' and 'conversation with colleagues'. The most important differences are that my classification is centred on the normative organisation of those 'encounters', and that it factorises the commitments present in 'encounters with colleagues' into two different types of entities: argumentative and practical commitments (and so, we could called the former 'encounters with language').

[14] This can also related to Kitcher's notion of a 'virtual' consensus; cf. Kitcher (1993), ch. 3, § 11.

[15] Jarvie (2001) offers a 'constitutional' interpretation of the role of methodological norms in the philosophy of Popper, particularly in connection with Sir Karl's defence of liberalism.

[16] Bicchieri (1988) discusses the properties an equilibrium in the choice of methodological norms must have to serve as a 'convention' in the sense of Lewis (1969).

[17] Lütge (2004) offers some examples of 'Prisoner Dilemma' games in the history of science. On the other hand, co-ordination problems can be exemplified by cases when different schools employ different units of measurement or technical standards.

[18] Perhaps we could interpret Feyerabend's 'anything goes', not as telling that 'all research strategies are equally functional under any circumstance', but as 'for every strategy, there are some conceivable circumstances where it is functional'. But it is also doubtful that this less stringent claim is actually true, and, even if it were, it would not be very useful, for it says nothing about how to identify the relevant circumstances.

[19] There is another sense of 'incommensurability', according to which the members of different scientific communities employ (or attach a very different meaning to) 'the same' (or 'similar') concepts. I think inferentialism can provide an interesting explanation of this phenomenon, one that hinders its purported relativistic consequences. The basic idea is that, by reducing meanings to patterns of inference, each community is able to identify the patterns followed by the other, and can make a comparison of the efficiency of each set of patterns according to its own goals.

Jesús P. Zamora Bonilla is professor in the Department of Logic, History, and Philosophy of Science of the National Open University of Spain (U.N.E.D.), Madrid, and academic co-ordinator of the Urrutia Elejalde Foundation for Economics and Philosophy. His research has centered on the economic analysis of scientific knowledge production, as well as the problem of scientific realism and truth approximation.