VERISIMILITUDE AND THE DYNAMICS OF SCIENTIFIC RESEARCH PROGRAMMES*

JESÚS P. ZAMORA BONILLA

SUMMARY. Some peculiarities of the evaluation of theories within scientific research programmes (SRPs) and of the assessing of rival SRPs are described assuming that scientists try to maximise an 'epistemic utility function' under economic and institutional constraints. Special attention is given to Lakatos' concepts of 'empirical progress' and 'theoretical progress'. A notion of 'empirical verisimilitude' is defended as an appropriate utility function. The neologism 'methodonomics' is applied to this kind of studies.

Key words: economics of science, empirical progress, epistemic utility, research programmes, scientific progress, theoretical progress, verisimilitude methodonomics

1. INTRODUCTION

After the work of Kuhn and Lakatos, there is little doubt that the process of scientific research is carried out on the basis of units which are larger than single theories. This fact, grounded basically on the evidence of history, has been a constant challenge for those philosophies of science which tried to explain the process of research as a rational activity in the epistemic sense. The most important problem seemed to be the reticence of scientists to abandon formerly accepted theories when these had been falsified by strong empirical evidence. This was clearly incompatible with the view of science as a pursuit of true knowledge about the world, both under the Popperian (hypothetico-deductive) and the Carnapian (inductivist) interpretations of scientific method. Neither Kuhnian 'paradigms' nor Lakatosian 'research programmes' were easy to handle either by classical falsificationism or by confirmationism, since these approaches were especially conceived to assess the epistemic value of single statements. The acceptance or rejection of scientific statements became, hence, a question where logic and empirical evidence had merely a partial role, perhaps less important than the role of psychic, social or cultural forces.

Journal for General Philosophy of Science **33:** 349–368, 2002.

A defensive strategy of some rationalist philosophers was to assume that scientists do not pursue directly 'true knowledge', but theories which, in spite of being strictly speaking false, are notwithstanding 'closer and closer to the truth'; *i.e.*, theories which are 'better and better' descriptions of the world. This strategy was initiated by Popper (1962) and followed by a number of authors especially during the seventies and the eighties.[1] Its basic rationale was that, even if a corpus of empirical evidence E has refuted two theories, A and B, it is still possible, under some interpretations of the concept of 'verisimilitude', that E *corroborates* the meta-hypothesis 'A is more verisimilar than B', rationally justifying so the 'preference' for A.[2]

The verisimilitude programme, however, suffers from a number of important problems, which undermine the possibility of using it to understand the actual development of science (see, for example, Zamora Bonilla, 1992 and 2000). As an alternative, I propose a 'methodological approach' to verisimilitude, one in which this concept is defined as an *epistemic* notion (rather than as a logical one), expressing the *perceived similarity* between a theory and the 'empirical truths' which a scientific community has established. Some interesting methodological conclusions derivable from such a definition have been offered in Zamora Bonilla (1996b). My aim in what follows is to show that this methodological approach to verisimilitude can explain also some streaking features of the process of theory assessment which takes place under the evolution of 'paradigms' or 'scientific research programmes' (SRPs). Of course, the model should only be taken as *an idealised representation of scientific activity*, and so, further 'concretizations' will be surely needed to approach still more this vision of science to the actual research practice.

In section 2 the core of my model of empirical verisimilitude is resumed; in section 3 I study the difference between 'assessing theories within an SRP' and 'assessing SRPs as such'; finally, in section 4 I consider the institutional framework of scientific activity as an unavoidable element to understand the development of SRPs from a 'methodonomic' approach, that is, one which sees scientists as rational decision makers trying to optimise some utility funcion under economic and institutional constraints.

2. A SIMPLE MODEL OF EMPIRICAL VERISIMILITUDE

I propose to understand the '*empirical verisimilitude*' of a scientific theory as a combination of two things: a) the degree in which it *seems closer* to the *empirical truths* which have been actually found out, and b) the degree of *information* or *content* of those empirical truths. This is, of course, an *epistemic* concept of verisimilitude (as opposed to a logical or semantical one), and it has also to be taken as a *relative* concept, in the sense that different scientific communities can assign different levels of empirical verisimilitude to the same theories on the basis of the same empirical facts. *Intersubjectivity* (as well as 'translation invariance') can be saved by showing that, accepting the same definition of verisimilitude, every scientific community should accept at least *the same methodological rules* (those derivable from that definition), and this acceptance will force at least *some* agreements on the evaluation of theories. I have offered elsewhere a defence of this pragmatic, epistemic interpretation of verisimilitude (Zamora Bonilla 1992 and 2000), and, for the sake of brevity, I will not repeat here those arguments.

A possible formalisation of this concept of *empirical verisimilitude* is the following:

(1) $Vs_1(T,E) = [p(T\&E)/p(T\vee E)]/p(E) = p(T,E)/p(T\vee E)$,

where T is a theory, E is a conjunction of empirical regularities, and $p(X)$ abbreviates '$p_c(X,B_c)$', that is, the probability assigned to X by the community C given the basic presuppositions or background knowledge of that community (B_c). So, $p(X)$ might be read as a 'subjective prior probability' in the Bayesian sense, though the precise interpretation of that function is not important now. My minimal assumption is that scientists in a research community can make at least *qualitative comparisons of probability* between different statements, and that these comparisons can be represented numerically by *any* assignment of probability values to individual statements which is consistent with the axioms of the probability calculus. Hence, my formal definitions of empirical truthlikeness can even be taken as mere analytical devices to derive *qualitative comparisons of verisimilitude*.

Another important comment concerns the nature of E. In real scientific practice, theories are usually not contrasted directly with singular data; rather, from those data an *empirical regularity* is somehow inferred, which is what researchers compare with their theories' predictions. In fact, a longer chain of inferences and comparisons is necessary, and what on one level is a hypothesis used to explain a set of data, becomes later a datum on a deeper level of explanation, though the same statement can not be a datum and a hypothesis in the same level.

Definition (1) leads to a set of reasonable methodological rules, basically derived from the fact that, if T is corroborated by E (i. e., if T entails E), then $Vs_1(T,E) = p(T)/p(E)^2$, and if T is confirmed by E (i. e., if E entails T), then $Vs_1(T,E) = 1/p(T)$. But it is certainly a 'naive' definition, at least because it does not take into account the fact that the empirical evidence is the combination of different empirical regularities, being each one usually asserted under a given margin of imprecision. So, if E_i represents any single regularity contained in E, if $K(E)$ is the set of all conjunctions of at least one E_i (plus the tautology), and if X_u is the statement asserting that X is true under the margin of imprecision u, then we can define:

(2) $Vs_2(T,E) = max\ (F \in K(E))\ Vs_1(T,F)$

(3) $Vs_3(T,E) = Vs_1(T, \cap E_{iu})$

(4) $Vs_4(T,E) = Vs_2(T, \cap E_{iu})$.

These definitions lead to almost the same methodological norms that (1), though they can give a positive degree of verisimilitude to theories that are falsified by E (*e.g.*, if they are not in contradiction with *all* the E_i's, or, in the last two cases, if they are 'close enough' to the E_i's). This possibility explains why we can take these functions as measures of empirical *verisimilitude*, rather than as measures of empirical *support*: falsified theories may sometimes have a high degree of empirical verisimilitude.

One of the most interesting methodological norms derivable from (2) is the following:

(5) Let E_1, E_2, ..., E_n represent n empirical regularities actually found out by a scientific community, and let T and S be two arbitrary theories. If:
i) any conjunction of the E_i's is statistically independent of any conjunction of the other E_i's;
ii) they can only be in one of the following three relations with any theory $U \in \{T, S\}$: U entails E_i, U is falsified by E_i, or U and E_i are statistically independent; if several E_i's are independent of U, then it is also assumed that U is independent of their conjunction;
iii) $p(T) < p(S)$, and
iv) every E_i entailed by T is entailed by S, and every E_i falsifying S also falsifies T,
then, $Vs_2(T,E) < Vs_2(S,E)$.

Proof: In the first place, suppose that all the E_i's contradict both T and S. In this case, since we have introduced the tautology into the set $K(E)$, the verisimilitude of both theories is equal to their prior probability. In the second place, if either T or S explain some E_i, we will first prove that, if *i* and *ii* are true, then $Vs_1(U,E) = Vs_1(U,C(E,U)\&I(E,U))$, where $C(E,U)$ and $I(E,U)$ represent respectively the conjunction of those laws in E which are derivable from a theory U, and the conjunction of those which are independent of U. Suppose that $Vs_2(U,E)$ is $Vs_1(U,F) > 0$, for some conjunction F of empirical facts contained in E, but not identical to E. Let E_i be an empirical fact contained in E but not in F. If E_i contradicts U, then $Vs_1(U,F\&E_i) = 0$; if E_i is statistically independent of U, then $Vs_1(U,F\&E_i) = p(U)/p(C(F,U)p(Uv(F\&E_i)))$; and if E_i is entailed by U, then $Vs_1(U,F\&E_i) = p(U)/p(C(F,U)p(E_i)p(Uv(F\&E_i)))$. It is easy to check that, in the last two cases, $Vs_1(U,F\&E_i)$ is higher than $Vs_1(U,F)$; hence, contrary to our assumption, this can not be the value of $Vs_2(U,E)$ if there are some E_i compatible of U and which is not contained in F. So, the conjunction of empirical facts which maximises the value of $Vs_2(U,E)$ is $C(E,U)\&I(E,U)$.

Now, the condition *iv* means that all the empirical facts which are explained by T are also explained by S, and that all the empirical facts which falsify S also falsify T. If this is true, then $C(E,S)\&I(E,S)$ will entail $C(E,T)\&I(E,T)$ (since the empirical facts which contradict T but not S will belong either to $C(E,S)$ or to $I(E,S)$). This, together with the assumption that S is more probable than T, entails that it will be also more verisimilar, according to the expression of Vs_2 given in the first part of the proof. *Q.E.D.*

Comparisons of theories allowed by (5) and by the other methodological theorems derivable from (1)-(4) are revisable as the empirical evidence changes. They can also be revised when what changes is our knowledge about the logical or statistical connections between theories and the empirical evidence.

Conditions *i* and *ii* put a limit to the applicability of (5), but they seem to be reasonable in many cases. Take into account that the E_i's are not whatever statements it is logically possible to construct, but empirical laws *actually found out* by a real scientific community. In the same way, T and S do not represent any pair of theories logically conceivable, but real theories proposed by real scientists, and there will be strong conceptual limits to the possible theories that it is rational to propose (*e.g.*, no Boolean combination of the empirical laws will be acceptable as an 'explanatory theory' of these same laws); within our conceptual framework, these limits will be included in the community's background knowledge.

For an example, consider Galilei's laws of falling bodies and Kepler's laws on planets as *explananda* of Newton's theory of gravitation. The acceptance of some of those laws did not force scientists to consider the rest 'more (or less) plausible', and so, it is reasonable to represent them as statistically independent from the point of view of the 'subjective probability functions' of XVIIth century scientists. Surely, in other cases some empirical findings will increase the probability of some new ones: *i.e.*, $p(E_{n+1}) < p(E_{n+1}, E_1 \& E_2 \& ... \& E_n)$; in the limit we may have that the second term of this inequality tends to 1, but, if this takes place, it seems rational to represent this fact by 'collapsing' all the E_i's into a single empirical law, E, which would be independent of the other laws (F, G, ...). Think, for example, of the E_i's as laws of the form 'the orbit of planet i is elliptical' (which is a universal statement referring to *all* the points in the orbit) and E as 'all planetary orbits are ellipses'.

Condition *iii* indicates that the conclusion '$Vs_2(T,E) < Vs_2(S,E)$' *might* not be true if S were less probable than T, in spite of *iv* being fulfilled. That is, a 'successful' theory can have a low empirical verisimilitude if it is extremely improbable under the basic assumptions of the community. Suppose that a theory S has been proposed whose core assumptions are contrary to the epoch's 'common sense' (and, so, has a low prior probability), but which correctly explains and predicts more empirical facts than the old theory T. Theorem (5) predicts that defenders of the old point of view will not necessarily accept the new theory; in fact, for them T may be still more verisimilar than S, in spite of their recognising the higher empirical success of S. Under these circumstances, the proponents of the new theory can try to increase its empirical verisimilitude by developing philosophical arguments (or scientific arguments of a 'higher level') making more credible their 'strange' hypotheses, *i.e.*, by increasing their 'prior' probability. Of course, they can only do that if they modify the accepted background assumptions of the community. This is what happens, at least according to Kuhn, during scientific revolutions. For example, geocentrist astronomers did not accept the verisimilitude of Copernicus' and Kepler's theories because these conflicted with some basic common sense assumptions of those days, though they might plainly accept that the new theories *did* explain the positions of planets *at least as well* as the geocentric theories did. During the first decades of the twentieth century, many scientists were reluctant to accept Einstein's ideas, as well as Quantum Mechanics, for similar reasons: they acknowledged the empirical success of these theories, but found their basic assumptions absurd. Keynesian economics explained the persistence of high unemployment rates, the instability of private investment and certain features of consumers' aggregate behaviour, and successfully predicted the positive effects of active government policies in solving economic crises, though it was not accepted by many economists because it was in conflict with certain core assumptions of classical economics relative to the notions of rationality and equilibrium. In geology, Wegener's theory of continental drift was also rejected during decades on similar bases: it explained more facts than its rivals, though it was inconsistent with the received ideas about the earth's physical nature.

A theorem close (though not identical) to (5) was Popper's desideratum when he introduced his concept of verisimilitude: what Popper unsuccessfully tried to show was that a theory *having passed more tests* than a second one was a (non-conclusive) reason to accept that the former has a higher verisimilitude.[3] Of course, what Popper would surely not accept is the limitation established in condition iii); his famous preference for 'bold' theories seems to imply that, given the other conditions, S would be better than T *even* if $p(S) < p(T)$, or *especially in that case*. I think, however, that what makes of 'boldness' a scientific virtue is the fact that, if a theory has to explain or to predict an *increasingly* stronger set of empirical regularities, it must have a low prior probability, since this has to be lower than the probability of the conjunction of all its empirical consequences. But if we fix our attention in a single moment of time, when a finite set of empirical regularities has been found, and when a limited number of theories have been proposed, then to have a low probability is not a virtue *per se*; for if it were, then it would be a 'child's play' to invent a better theory: just add to one successful theory whatever conjunction of statements consistent with it; the resulting theory will have at least the same empirical consequences of the previous one, but its prior probability will be much lower. It can be argued that this new 'theory' would be *ad hoc*, and hence unacceptable from the falsificationist point of view; but its adhocness refers in this case to *its incapacity to offer new predictions*, not to its degree of prior probability or informativeness.

Moreover, there is a reason –and a strong one indeed!– for inventing 'bold' theories: definitions (1)-(4) entail that the *maximum* degree of truthlikeness a given theory T can reach is $1/p(T)$ (when E entails T); so, launching a new, improbable hypothesis can sometimes be the only way of increasing the empirical verisimilitude of scientific knowledge. In the same way, the maximum verisimilitude that a set E of empirical regularities can give to any theory is $1/p(E)$ (this takes place when T is logically equivalent to E), what offers also a reason to find out new empirical regularities.

3. EVALUATION OF THEORIES WITHIN SCIENTIFIC RESEARCH PROGRAMMES

A SRP is usually defined as a *set of theories* having some hypotheses in common (the 'hard core') because a scientific community has decided not to eliminate them in spite of the falsification of any of those theories by the empirical evidence. In a pragmatic sense, however, the SRP is defined by the community's *intention* of 'fitting the world' within the hypotheses of the core, finding a set of 'auxiliary hypotheses' (the 'belt') which, joined with the core, shows the highest scientific value, whatever defined (see Lakatos, 1978). The set of theories constituting the SRP is by no means fixed, and it grows through time, as new alternative hypotheses or 'protective belts' are invented and as the SRP is applied to new sets of empirical events. The last fact usually gives the SRP the form of a *net*, with wider sets of empirical applications to which only the 'core' and some very general hypotheses are applied, and more limited subsets of applications that the SRP tries to explain by attaching to the former hypotheses some more specific ones (see Balzer, Moulines and Sneed 1987).

In each level of such a net there are usually several *alternative* theories. That is, the theories in an SRP do not only compete with the theories of *other* SRPs: even in the absence of rival programmes, there is always more than one theory, and scientists may want to know which of them is 'the best'. Criteria to compare *theories within an SRP* are, hence, as necessary as criteria to compare *competing SRPs*.

3.1. *Assessing Theories Within an SRP (I): Empirical Progress*

Popular expositions of Lakatos' methodology assert that the different 'protective belts' of each SRP are usually developed according to some internal logic or preconceived plan (the 'positive heuristic'). But this plan only seems to have a definite role during the very first steps of the programme's development; in the later stages of SRPs their evolution is probably more inspired by the discovery of new empirical regularities. In fact, this discovery is constantly motivated by SRPs themselves, since findings are the product of former attempts of testing the SRP's theories. It is the unforeseeable evolution of the empirical evidence (and hence, the world, answering our questions formulated in controlled experiments or observations), combined with the imagination and talent of scientists, what ultimately 'explains' the development of most SRPs.

In this context of accumulating empirical evidence, and especially during the first moments of an SRP's evolution (when few empirical data can still be taken into account), it is reasonable to assume that researchers are not just interested in the *present* value of their theories' empirical verisimilitude: since the empirical evidence is expected to change at a high rate, a theory which has a low verisimilitude according to the *known* empirical laws might perhaps be corroborated by many *future* findings, and so, its verisimilitude may grow, and *vice versa*. So, in these cases it seems more rational to evaluate theories on the basis of the *expectations* of researchers about the possible future scenarios of empirical knowledge. Within the frame of our model, this can be understood as if scientists were interested in the *expected verisimilitude* of theories.[4] This expected value can be expressed as follows:

$$(6)\quad Vs_5(T,E) = \Sigma\ (x \vdash E)\ p(x,E)Vs_1(T,x)$$
$$= \Sigma\ (x \vdash E\&\neg T)\ p(x,E)Vs_1(T,x) + \Sigma\ (x \vdash E\&T)\ p(x,E)Vs_1(T,x)$$
$$= 0 + \Sigma\ (x \vdash E\&T)\ p(x,E)p(T,x)/p(Tvx)$$
$$= \Sigma\ (x \mid\!\!- E\&T)[p(x)/(p(E)][1/p(T)]$$
$$= p(E\&T)/(p(E)p(T))$$
$$= p(E,T)/p(E)\ \text{(A)}$$
$$= p(T,E)/p(T)\ \text{(B)}^5$$

where the x's are the minimal points in which the logical space can be divided (for example, logical constituents).

(6.A) and (6.B) allow to derive two interesting methodological norms, respectively:

(7) In any moment of time, the theory with the maximum expected verisimilitude will be that with the maximum likelihood (*i. e.*, $p(E,T)$).

(8) The expected verisimilitude of a theory increases or decreases simultaneously with its empirical support or degree of confirmation (*i.e.*, with $p(T,E)$).

As we have seen, theories within an SRP do not form necessarily a one-dimensional series; usually, several incompatible 'protective belts' are developed simultaneously for the same 'core' in the same application, but even when only one 'belt' is developed each time, the community has to choose the best one, at least to know whether the SRP has 'progressed' or 'degenerated'. To make this choice, (7) tells that, if researchers rationally pursue a high expected verisimilitude, their preferred criterion should be that of maximum likelihood; that is, the theory with the highest *explanatory success* should be preferred.

On the other hand, (8) states that the evolution of a single theory's expected verisimilitude is identical to the evolution of its *degree of confirmation*.

All this entails that there is a tension between the features a good theory must have, if the epistemic attitudes of researchers were described by Vs_5. To be better than the other theories, it has *to explain as many facts as possible*, and so, it has to be a 'strong' or 'bold' theory (as Popper desired); but to become still better, it has to be *strongly confirmed by empirical data* (as Carnap insisted).

From Vs_5 we can derive more complicated definitions of verisimilitude, analogous to (2)–(4):

$$(9)\qquad Vs_6(T,E) = max\ (F \in K(E))\ Vs_5(T,F)$$

(10) $Vs_7(T,E) = Vs_5(T, \cap E_{iu})$

(11) $Vs_8(T,E) = Vs_6(T, \cap E_{iu})$.

Now, the analogue of theorem (5) is the following:

(12) Let $E_1, E_2, ..., E_n$, T and S be as in (5). If:

i) any conjunction of the E_i's is statistically independent of any conjunction of the other E_i's;

ii) they can only be in one of the following three relations with any theory $U \in \{T, S\}$: U entails E_i, U is falsified by E_i, or U and E_i are statistically independent; if several E_i's are independent of U, then it is also assumed that U is independent of their conjunction; and

iii) every E entailed by T is entailed by S,

then, $Vs_6(T,E) < Vs_6(S,E)$.

Proof: An argument similar to that of (5) shows that, if *i* and *ii* are valid, then $Vs_6(U,E) = 1/p(C(E,U))$ (or 1, if all the elements in $K(E)$, save the tautology, falsify U). On the other hand, if *iii* is true, $p(C(E,T))$ will be at least as high as $p(C(E,S))$, and so $Vs_6(S,E)$ will be at least as high as $Vs_6(T,E)$. *Q.E.D.*

This theorem asserts that, if a theory explains at least all the empirical laws explained by its rivals, then it will be at least as good as them from the point of view of Vs_6, *independently of the cases where it is falsified*. For example, if there are five empirical laws, T entails E_1 and is not refuted nor corroborated by the other laws, and S entails E_1, E_2, E_3 and E_4, but is falsified by E_5, then S will be better than T in spite of the fact that S, contrary to T, has been empirically refuted. This seems to be consistent with Lakatos' indication that the falsification of the 'first' models developed in an SRP does not worry very much to their creators; they only become to take falsifications seriously (that is, they begin to use (5) instead of (12)) when new strong empirical findings are not deliberately produced nor expected, that is, when SRPs have been fully developed to take into account a more or less fixed *corpus* of empirical regularities. It is only in that moment when 'unsolved problems' become 'anomalies' for an SRP.

Note also that we have not in (12) the condition '$p(T) < p(S)$'. This can be interpreted as telling that, during the 'hot' moments of an SRP's development, no limits are put to the researchers' imagination, and each hypothesis is only valued according to its explanatory or predictive success, not by its prior plausibility. This last criterion begins only to be applied when rival programmes are 'mature' and only 'marginal' additions to the empirical knowledge are expected. Lakatos refers, for example, to Bohr's electron theory and its inconsistency with Maxwellian electrodynamics; Bohr's theory had zero probability according to the background knowledge, but this was not seen as an obstacle to develop it as long *as its empirical success was growing*.

3.2. *Assessing Theories within an SRP (II): Theoretical Progress*

The standard mathematical device employed in (6) is also applicable to different kinds of situations; for example, it can be applied when *new empirical predictions* have been made but have not been actually tested still (*e.g.*, because of limited resources). The comparison of theories taking into account these new predictions could run as follows. Assume that each prediction has the form 'under conditions C, result A obtains', and that the predictions of any two theories about what happens under C (if both make such a prediction) are identical or mutually contradictory; let E be the conjunction of all empirical regularities already found; let F_i be the conjunction of one *formulated* prediction for each considered situation C (including the statements which affirm that no existing prediction is true about C), and let $Q(F)$ represent the set of all the F_i's. Assume also the hypotheses *i* and *ii* of (5) and (12) for all the combinations of known regularities and not tested predictions. In this case we will have the following definition of the '*hypothetical verisimilitude* of theory T given the empirical evidence E and the set of predictions $Q(F)$' ($N(F_i,T)$ represents the conjunction of those laws in F_i which are not explained by T, because it gives a different prediction or because it gives none):

(13) $$HVs_6(T,E,Q(F)) = \Sigma\ (F_i \in Q(F))\ p(F_i,E)Vs_6(T,E\&F_i)$$

$$= \Sigma\ (F_i \in Q(F))\ p(F_i)[1/p(C(E\&F_i,T)]$$
$$= \Sigma\ (F_i \in Q(F))\ [p(C(F_i,T))p(N(F_i,T))]/[p(C(E,T))p(C(F_i,T))]$$
$$= [\Sigma\ (F_i \in Q(F))\ p(N(F_i,T))]/p(C(E,T))$$

(Note: if T entails all the predictions included in one F_i, then F_i is identical to $C(F_i,T)$, and hence the term '$p(N(F_i,T))$' in the last line has to be replaced by 1 for that F_i).

I have used Vs_6 instead of Vs_2 because in this case it is reasonable to assume also the uncertainty about the result of future empirical tests. Of course, any other definition of verisimilitude or epistemic value can be used.

The analogue of (5) and (12) derivable from (13) is:

(14) Let $E_1, E_2, ..., E_n$, T and S be as in (5), and let $Q(F)$ be as in (13). If:

i) any conjunction of empirical results (tested or not tested) is statistically independent of any conjunction of the others;

ii) each empirical result can only be in one of the following relations with any theory $U \in \{T, S\}$: U entails it, or U is falsified by it, or U is statistically independent of it; if several empirical results are independent of U, then it is also assumed that U is independent of their conjunction;

iii) T and S have the same logical connections with each E_i; and

iv) all predictions made by T are also made by S, and there is some prediction made by S about which T does not tell anything,

then $HVs_6(T,E,Q(F)) < HVs_6(S,E,Q(F))$.

Proof: Let F_i be the statement $A_1\&A_2\&...\&A_m$; we can order the A_i's in such a way that T entails A_1 to A_k ($k < m$); from *iv*, S will entail besides at least some A_j ($j > k$); hence, for each F_i, we have that $p(N(F_i,T)) = p(A_{k+1}\&...\&A_j\&...\&A_m) < p(A_{k+1}\&...\&A_{j-1}\&A_{j+1}\&...\&A_m)) < p(N(F_i,S))$ (the last inequality becomes a strict one if, besides A_j, S makes other predictions which T does not make). The proof is completed taking into account that *iii* entails that $p(C(E,T)) = p(C(E,S))$. *Q.E.D.*

This situation can be seen as a case of *theoretical progress* in Lakatos' sense, that is, as a reason to consider that a theory is 'better' than another just because the first one makes more predictions than the second one, but before those predictions are effectively tested.

3.3. *Comparing SRPs*

The question about the comparability of SRPs has a very simple answer within the model developed in this paper: to compare two competing SRPs by their empirical verisimilitude is to compare the most verisimilar theories of each one.[6] That is,

(15) $Vs(P,E) < Vs(P',E)$ if and only if there is a theory T' in programme P' such that, for every T in programme P, $Vs(T,E) < Vs(T',E)$.

All previous definitions of verisimilitude are applicable to (15), and, of course, they can be changed by other 'epistemic utilities'. I think that the same selection criteria can now be used: in contexts of strong competition and rapid growth of the empirical evidence, Vs_6 or Vs_8 will be preferred; in contexts where empirical research is delayed with respect to a feverish theoretical activity, the sensible options are HVs_6 or its possible derivative HVs_8; and, finally, when the scientific community feels that all the major empirical and model-making work with competing SRPs has been already carried out, Vs_2 and Vs_4 will be used to compare among them.

The main difference between the comparison of theories and the comparison of SRPs is that, in the former case, comparisons of empirical verisimilitude are supposed to be identical to comparisons of *preference* among theories, while this identity ceases to be valid when we compare SRPs. That is, when we assert that $Vs(T,E) < Vs(S,E)$, we are asserting that, in the light of E, S seems at least as good as T, independently of whether T and S belong to the same SRP or not. But, in any moment of time, no theory actually contained in a programme can be taken as 'the last word': it is always possible that defenders of a 'regressive' SRP find a 'protective belt' that makes the new theory more verisimilar than all theories in the rival programmes. This means that an SRP can not be rejected by considerations of empirical verisimilitude alone (or, in general, by epistemological considerations), while individual theories within one SRP can. In the next section we will see that the reasons to abandon an SRP are, rather, *institutional* ones (but not 'irrational'!).

In any case, it seems reasonable to suppose that, within a single SRP, those lines of development or 'subprogrammes' which are *more plausible* from the point of view of the hard core will be explored first. This means that, if a previous version (say $C\&B$, where 'C' is the core and 'B' is one possible belt) of the SRP has many anomalies, researchers can propose a new version $C\&B'$, which perhaps explains some of the anomalies, but at the cost of lowering the prior probability of the theory (since $p(B',C) < p(B,C)$, we have that $p(C\&B') < p(C\&B)$). We can interpret this as an example of *adhocness* (in the sense that the second theory seems to be '*ad hoc*', since it has been developed in order to avoid a certain anomaly of the first theory), and as one possible reason (among others) to explain *why it is better for an SRP to predict a new fact than to explain an old one*: assuming that the context makes more reasonable to use (5) instead of (12) or (14), had those facts been correctly predicted by $C\&B$, the degree of empirical verisimilitude of *that* theory would have been higher than that of $C\&B'$, even if the latter explains them perfectly, and hence, the verisimilitude of the SRP *could have been* also higher.

We have an example of this in the case of geocentric astronomy: to save the programme's basic hard core assumption (the earth's rest), the first simple hypotheses of regular and circular planetary motions had to be replaced by increasingly complicated and anti-intuitive ones, which gave a growing implausibility to the programme.

4. INSTITUTIONAL ASPECTS OF SCIENTIFIC RESEARCH PROGRAMMES: A METHODONOMIC APPROACH

The development of an SRP is a job to which many of our most intelligent fellow citizens devote almost their full lives. Scientific activity involves a tremendous expenditure in human, financial and technological resources, and so, it would make no sense to believe that the social, economic and institutional framework of scientific activity has no noticeable effect on the content of scientific knowledge nor on the materialisation of scientific practices. They can influence science as epistemic reasons do, but, instead of supposing that 'society' and 'epistemology' are conflicting *explanantia* of scientific activity, we can try to unify them through the hypothesis that scientific researchers are a kind of *economic agents* who pursue a utility function with epistemic variables among the most relevant ones, and whose decisions are constrained by economic, social, cultural and psychological limits. I propose to call '*methodonomics*' this kind of approach to scientists decision making.[7] In this last section I shall briefly apply this approach to the study of two particular questions: why is scientific research usually carried out through SRPs instead as through isolated theories?, and how are resources distributed among the different participants in an SRP and among the different tasks they have to perform?

4.1. *Why are there SRPs?*

The idea that the units of scientific activity are not single theories but larger sets of them was offered during the 1960's to accommodate falsificationist methodology to the growing historical evidence against its most naive versions. More specifically, it was an attempt to explain why the main falsificationist rule ('reject any theory which has been contradicted by empirical evidence') was not apparently followed in many cases and by many great scientists. In other words, the articulation of scientific activity by means of SRPs was originally an *explanans* of a historic datum, as in the following schema:

Explanandum-1: Why do scientists continue accepting theories that are in conflict with experience?

Explanans-1: Because scientists try to evaluate programmes, not theories.

Besides the illuminating comments of Kuhn (1962), I do not know many works devoted to a question of a higher level, which is to take our *explanans*-1 as a new *explanandum*:

Explanandum-2: Why is scientific activity structured in SRPs and not in individual theories?

My hypothesis (which can be taken as an economic reading of Kuhn, rather than a 'psychological' or 'social' one) is that the use of SRPs is a way to maximise the field of application of scientific knowledge, minimising the cost of inventing theories and of working with the conceptual apparatus associated to them. In the terms of the preceding section, scientists try to maximise the empirical verisimilitude of their theories, what requires to explain a large set of empirical regularities, but what is costly in terms of intellectual effort (at least). Creating a new scientific idea involves usually a great deal of imagination and a big amount of time to the most talented researchers of a discipline. The development of this new idea requires also to learn sophisticated techniques (mathematical, experimental or both); and, as Kuhn noted, their acquisition by students makes that the assimilation of alternative ideas can be an almost unfeasible task in the future.

So, if scientific work were to be carried out by means of independent theories (that is, if each theory had a set of conceptual and formal instruments different from the others'), then too much time and effort should be devoted to the creation and implementation of those theories, and too little time and effort would be left to the application of theories to new empirical cases. Exaggerating a little, we could say that, had the Kuhnian 'normal scientist' unlimited intellectual and economic resources, he would become immediately a Popperian falsificationist.

This also explains why SRPs are usually organised under the form of *nets*: when a group of researchers are interested in a concrete application, to develop a completely original theory is much more expensive than to go to a general field, to assume the general laws accepted in that field, and to try to invent a particular law that, joined to the general ones, produces a theory empirically adequate to their concrete application. The iteration of this strategy, partly promoted by the social pressure to quickly develop strong empirical results in well defined fields, will tend to produce the classical net-like structure of most SRPs.

Using another economic metaphor, we could say that the creation of an SRP creates 'economies of scale'; it is a large 'initial investment' with a high cost at the beginning, but, as it serves to produce a growing number of different results, it makes that the 'unit cost' of each one of its uses is diminishing.

But, if all this is true, why then is there *more than just one* SRP in each discipline? If it is costly to develop alternative ideas, why are they developed sometimes anyway? This objection to the argument above forgets that science does not only try to minimise its costs, but basically to maximise its results: 'investment' in the development of alternative programmes may be 'profitable' if the actual SRP leaves unexplained many empirical data. In a case like this, the scientific community – or some part of it – will perceive a possibility of reaching a higher degree of verisimilitude by trying to develop a new SRP than by devoting its effort to further implementing the old programme.

From the point of view of the individual researcher, if we assume that *her* 'utility' depends upon the empirical verisimilitude of the theories *invented by her*, she should guess the value of her 'expected utility' when following different courses of action: in some cases (if there are many 'anomalies' in the old SRP, and she feels she has 'promising new ideas') it will give more expected utility to take the risk of leaving orthodoxy; in some other cases (when the old SRP enjoys good health and the individual scientist is more 'risk averse') it will give more expected utility to swim with the stream.

Finally, it should be indicated that the very concept of an SRP is considerably fuzzy. In fact, *in every point of a theory-net can a sub-programme be initiated*, with a particular hypothesis as its 'hard core' together with the net's general laws. Since the cost of developing new ideas is surely higher when they affect to the full building of a discipline than when they only affect to a small part of it, and since the probability of finding out new successful ideas is higher in a little subfield than in the core of the discipline (which was so hard to learn!), it seems reasonable to expect that *the proliferation of rival SRPs will be higher in the low elements of large theory-nets than in its nuclear ones*. This can be a reasonable explanation of the debate between Kuhn and Lakatos about whether the history of science was a series of 'normal' periods with just one predominating paradigm each time, or a constant struggle between different research programmes.

4.2. *Allocating Scientific Resources*

Not only is the *existence* of SRPs influenced by the institutional features of the scientific communities where they are developed; also their *evolution* can be in part explained by those social aspects. For example, we have seen that the set of applications in which an SRP evolves depends in part on which fields have a 'social interest'. In other cases, the social pressure may be not so direct, but a different important reason for the selection of fields of application is also of social nature, though in this case its social character refers to the *institutional structure* of scientific communities: *competition* between individual scientists or teams may lead them to choose applications that (they think) favour their SRPs over their rivals' ones; or an application may be interesting because it serves to easily confront two or more SRPs. In the absence of rivalry, those applications could have not been studied forever. Rivalry and competition, then, are forces that influence the development of SRPs, but they can not be understood at all without referring to the scientists' interests and motivations.

Other aspects of SRPs that can be explained in a methodonomic fashion are the following ones:

a) Almost every resource can be employed in *alternative uses*. For example, money can be expended in buying more instruments or in engaging more researchers. Time and effort can be spent in theoretical or in experimental work. A journal can not accept all the articles subscribed to it (its pages are a limited resource), and so on. *Different institutional rules will lead to different decisions in each one of these cases*. If decisions are taken by a central authority, perhaps scientific teams end up with lots of computers but too many few people to work with them. If each individual chooses his own field, the 'allocation of human capital' may be inefficient from the collective point of view. If referees were to be paid (for example, according to the future success of the papers they recommend to accept), perhaps different papers would be written and published, etc. So, *the rules of decision-making within the scientific community can have an effect on the results of scientific activity*. An interesting field of work in philosophy of science would be, hence, the study of this influence of institutional aspects on scientific performance.

b) A concrete example of decision-making is that of *choosing which experiments or observations to make*, when only some planned ones are allowed by the scientists' budget. From the ideas of section 3, it seems rational that scientists will base their decision in the utility they expect to reach under the assumption that certain experiments are made. Imagine a very simple case where there are two rival theories, *T* and *S*, which explain the present empirical evidence *E*. Two experiments have been conceived, *F* and *G*, but there is money for just one of them; *FT*, *FS*, *GT* and *GS* are the respective predictions of *T* and *S* about those experiments (for simplicity, I assume that they are the only possible results). Which experiment will be chosen? If the scientific community takes a *centralised* decision, it will choose the experiment which corresponds to the larger of the following two quantities:

$$(16) \qquad \begin{aligned} XVs(E,F) &= p(FT,E)Vs(T,E\&FT) + p(FS,E)Vs(S,E\&FS) \\ XVs(E,G) &= p(GT,E)Vs(T,E\&GT) + p(GS,E)Vs(S,E\&GS) \end{aligned}$$

This means that for the scientific community it is indifferent *which* theory 'wins'. On the other hand, if the decision is taken by someone who has an interest in a concrete theory (for example, *T*), he will choose the experiment that gives the larger of these two values:

$$U(T,E,F) = p(FT,E)Vs(T,E\&FT) + p(FS,E)Vs(T,E\&FS)$$
$$U(T,E,G) = p(GT,E)Vs(T,E\&GT) + p(GS,E)Vs(T,E\&GS) \tag{17}$$

That is, he will choose that experiment which gives a higher expectation of *T* becoming better than *S*. It is by no means necessary that both decision rules (the 'centralised' and the 'decentralised' ones) lead to making the same experiment, especially when the idealised situation assumed above is progressively complicated. So, *institutional rules* (in this example, the rules about who decides which experiments are to be made) *can affect the evolution of empirical evidence*.

c) Lastly, as Lakatos and Feyerabend noted, there can be *no definitive epistemic reasons that force to reject an SRP*. In fact, it is more difficult to explicate what is to 'accept' or 'reject' an SRP than to explicate what is to 'accept' or 'reject' a single theory. I shall not try such an explication now. Since the target of methodonomics is to explain scientists' *actions*, I shall consider that to accept an SRP is equivalent to publicly defend it against its rivals, or to devote time and effort to elaborate theories within it, or to derive predictions from them, or to test them, and so on. Under this assumption, to abandon an SRP is just to stop working with it. Hence, the obvious reason to abandon an SRP is that researchers have thought it gives more expected utility to devote their resources to other jobs. Depending on the institutional features of scientific communities, and hence, on the rules of distribution of 'merit' among them, different degrees of utility can be expected by each individual researcher or team if they choose to work with some SRPs or with others. This means that the 'life' of an SRP depends on the *perceived* probability of 'doing something remarkable' with it, and this will depend on the 'advantage' that the other SRPs have over the former in terms of the functions of verisimilitude studied in the previous sections.

NOTES

* Research for this paper has been funded by the Spanish Government's DGICYT, as part of research projects PB 95-0125-C06 and PB98-0495-C08-01. Some ideas were presented at a symposion on structuralist theory of science, held in Zacatecas (Mexico) on february 1998; I want to express my gratitude to its organisers and to the other participants in the symposion for helpful and illuminating comments on those ideas.

[1] See especially Kuipers (1987) and Niiniluoto (1984) and (1987). For the Spanish reader, Zamora Bonilla (1996a) offers a detailed history of the verisimilitude programme.

[2] This point is made especially clear by Kuipers (1991) and (1996). Kieseppä (1996) has criticised the attempt of giving some methodological force to the theory of verisimilitude. I think that what follows in the present paper can be taken as a counterexample to his arguments, which I have analysed more deeply in Zamora Bonilla (2000).
[3] See Popper (1963), and the articles of Kuipers referred to in note 2.
[4] This 'expected verisimilitude' is not identical to Niiniluoto's 'estimated verisimilitude' (see Niiniluoto (1987), p. 269), though both concepts employ the same mathematical device and are, of course, very close in their spirit. The main difference between Niiniluoto's concept and mine is that the former is an empirical estimation or 'measure' of an *objective* but unobservable magnitude (the theory's 'real' closeness to the truth), while my definition represents the expected value of a *subjective* preference. Of course, Niiniluoto's concept can also be read as the expected value of an 'epistemic utility' (*op. cit.*, ch. 12).
[5] Vs_5 is obviously equivalent to the statistical notion of 'multiplicative support' or 'ratio degree of confirmation', which is usually taken by many Bayesians as a measure of the acceptability of a hypothesis. See, for example, Howson and Urbach (1989), pp. 86 and ff.
[6] Comments about the verisimilitude of theory nets in Zamora Bonilla (1996b), sections 4 and 5, are also applicable now, though the present version is considerably simpler.
[7] Though 'methodonomics' is, as far as I know, a neologism, the subject is not my invention, of course; one important contribution is Kitcher (1993), whose ch. 8 develops a series of interesting 'methodonomic' models. From a different point of view, in Zamora Bonilla (1999) I have offered an economic model of the dynamics of theory acceptance, which is independent of the problem of the relation between SRPs and individual theories.

REFERENCES

Balzer, W., Moulines, C. U. and Sneed, J.: 1987, *An Architectonic for Science*, Reidel, Dordrecht.

Howson, C. and Urbach, P.: 1989, *Scientific Reasoning. The Bayesian Approach*, Open Court, La Salle (Ill.).

Kieseppä, I. A.: 1996, 'On the Aim of the Theory of Verisimilitude', *Synthese* **107**, 421–438.

Kitcher, P.: 1993, *The Advancement of Science*, Oxford University Press, New York.

Kuhn, T. S.: 1962, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.

Kuipers, T. (ed.): 1987, *What is closer-to-the-truth?*, Rodopi, Amsterdam.

Kuipers, T.: 1991, 'Revisiting the Hypothetico-Deductive Method', *Acta Philosophica Groningana* **3**.

Lakatos, I.: 1978, 'Falsification and the Methodology of Scientific Research Programmes', in I. Lakatos, *The Methodology of Scientific Research Programmes. Philosophical Papers. Vol. 1*, Cambridge University Press, Cambridge.

Niiniluoto, I.: 1984, *Is Science Progressive?*, Reidel, Dordrecht.

Niiniluoto, I.: 1987, *Truthlikeness*, Reidel, Dordrecht.

Popper, K. R.: 1962, 'Truth, Rationality and the Development of Knowledge', in *Conjectures and Refutations*, Routledge, London.

Zamora Bonilla, J. P.: 1992, 'Truthlikeness Without Truth: A Methodological Approach', *Synthese* **93**, 343–372.

Zamora Bonilla, J. P.: 1996a, *Mentiras a medias* (in Spanish), Ediciones de la Universidad Autónoma de Madrid, Madrid.

Zamora Bonilla, J. P.: 1996b, 'Verisimilitude, Structuralism and Scientific Progress', *Erkenntnis* **44**, 25–47.
Zamora Bonilla, J. P.: 1999, 'The Elementary Economics of Scientific Consensus', *Theoria* **14.3**, 461–88.
Zamora Bonilla, J. P.: 2000, 'Truthlikeness, Rationality and Scientific Method', *Synthese* **122**, 321–35.

Departamento de Lógica y Filosofía de la Ciencia
Facultad de Filosofía
U.N.E.D.
28040 Madrid (Spain)
(jpzb@urrutiaelejalde.org)