FACULTAD DE CIENCIAS ECONÓMICAS
Y EMPRESARIALES

SOME REFLECTIONS ON POPPER´S APPROACH TO RATIONALITY AND ITS IMPLICATIONS FOR THE SOCIAL SCIENCES

Iván H. Ayala
Alfonso Palacio-Vera

Working Papers / Documentos de Trabajo. ISSN: 2255-5471
DT CCEE-1503 Marzo 2015
http://eprints.ucm.es/29352/

SOME REFLECTIONS ON POPPER´S APPROACH TO RATIONALITY AND ITS IMPLICATIONS FOR THE SOCIAL SCIENCES

Abstract:
There are two different notions of human rationality in Popper´s work: the notion that stems from his evolutionary theory of knowledge and learning (PTKL), and the notion embodied in his methodological proposal for the social sciences known as `Situational Analysis´ (*SA*). This essay provides an in-depth critical analysis of the relation between these two approaches and its implications for the social sciences. In particular, we focus on (i) the difference between PTKL and both the `objectivist´ and `subjectivist´ version of *SA*, (ii) Schumpeter´s distinction between the `rationality of the observer´ and the `rationality in the observed´, and (iii) Hayek´s arguments about the nature of the `facts´ of the social sciences.

Keywords: Evolutionary, Knowledge, Popper, Rationality, Situational analysis.

ALGUNAS REFLEXIONES ACERCA DEL ENFOQUE POPPERIANO SOBRE LA RACIONALIDAD Y SUS IMPLICACIONES PARA LAS CIENCIAS SOCIALES

Resumen:
Existen dos nociones distintas sobre la racionalidad humana en la obra filosófica de Popper: la derivada de la teoría popperiana del conocimiento y el aprendizaje (TPCA) y la noción implícita en su propuesta metodológica para las Ciencias Sociales conocida como "Análisis Situacional" (AS). Este ensayo realiza un análisis crítico detallado de la relación entre estas dos nociones y de sus implicaciones para las ciencias sociales. En concreto, el ensayo aborda: (i) las diferencias entre TPCA y las versiones "objetivista" y "subjetivista" de AS, (ii) la distinción propuesta por Schumpeter entre la "racionalidad del observador y la "racionalidad del observado" y (iii) los argumentos de Hayek sobre la naturaleza de los "hechos" en las Ciencias Sociales. .

Palabras clave: Evolucionista, Conocimiento, Popper, Racionalidad, Análisis situacional.

Materia: Ciencias Sociales
JEL: A12, B41, B52, D81.

Iván H. Ayala
Universidad Complutense de Madrid
E-mail: ivanhuascarayala@ucm.es

Alfonso Palacio-Vera
Departamento de Economía Aplicada III (Política Económica)
Facultad de Ciencias Económicas y Empresariales. Universidad Complutense de Madrid,
E-mail: apv@ccee.ucm.es

Marzo 2015 (fecha de recepción)

SOME REFLECTIONS ON POPPER´S APPROACH TO RATIONALITY AND ITS IMPLICATIONS FOR THE SOCIAL SCIENCES

1. Introduction

Discussions by both philosophers of science and social science methodologists on Popper´s methodological proposal for the social sciences known as `Situational Analysis´ (*SA*) have focused either on its (in)compatibility with falsification (Hands, 1985, 1991, 1992; Notturno, 1998; Hedström *et al*., 1998) or on its interpretation.[1] In the former case, the debate has revolved around Popper´s confession (Popper, 1994, ch. 8) that his `rationality principle´ (*RP*), i.e., the notion that in the construction of models in the social sciences we must assume that actors´ behaviour is adequate or appropriate to their problem-situation (P-S), is *false* but nevertheless a *good enough approximation to the truth* (Popper, 1985).[2] This surprising confession by a philosopher of science whose academic reputation grew out of the formulation of a demarcation criterion for scientific hypotheses based on the requirement that the latter be potentially falsified has led some commentators to argue that Popper´s methodological proposal for the social sciences and, specifically, his *RP* is incompatible with the criterion of refutability as prescribed for the natural sciences.[3] Some critics even argue that such incompatibility severely undermines Popper´s claim to methodological monism in the natural and the social sciences. Other critics have accused him of reintroducing a pure instrumentalism *à la* Friedman (1953) due to his methodological advice to immunize *RP* from potential refutation and of `contradicting his own explicit rejection of instrumentalist pretences to knowledge´ (Oakley, 1999, p. 32).

[1] General criticisms of Popper´s methodological proposal for the social sciences can be found in Latsis (1983), Hands, (1985), Bunge (1996), Oakley (1999, 2002), and Vanberg (2002).

[2] Recently, another version of *RP* has been proposed in Lagueux (2006, pp. 201-202) who suggests that, given the fact that refinements in model-construction in the social sciences imply that theoretical models exhibit more detailed descriptions of the situation, *RP* may also be enunciated as implying that `the agent will agree with what is clearly presented by the model itself as the appropriate thing to do´. Be that as it may, for the purpose of this study we focus hereafter on Popper´s version of *RP*.

[3] These ideas were expounded in his famous *Logic of Scientific Discovery*. That said, and as far as the social sciences are concerned, there is some consensus on the notion that it is not 'falsificationism' *per se* but 'critical rationalism' — of which falsification is only one possibility — that is the true message of Popper´s philosophy (Caldwell, 1991; Notturno, 1998; Boland, 2003*a*). According to Hands (1991, p. 114), who nevertheless remains sceptical of this interpretation, `critical rationalism´ is an interpretation of Popperian philosophy due primarily to Bartley (1982).

As for the interpretation of *RP*, the debate has centred on its role and status. For instance, Latsis (1983, p. 132) argues that `Popper´s account of the role and status of the rationality principle is obscure and unsatisfactory´. He shows that, in different parts of his work, Popper notes that *RP* is `almost empty´, `not a priori valid´, `clearly false´, `a good approximation to the truth´, and `the consequence of a methodological postulate´ (*op. cit*., p. 133). According to him, the role of *RP* is to function as a `plastic interface´ between mental states and behaviour and this is the reason why it is declared by Popper to be false but close to the truth (*op. cit*., p. 140). Specifically, *RP* is false if interpreted in a literal way because it does not *determine* behaviour in a `cast-iron´ fashion but is close to the truth because it captures the tendency of human behaviour to follow the logic dictated by P-S.[4] Crucially, he distinguishes between an `objectivist´ (*RPo*) and a `subjectivist´ version (*RPs*), and associates the former with Pareto (1917, section 150) and Parsons (1937, p. 58), and the latter with Popper himself. If *RPo* is adopted, the theoretician reconstructs the `objective´ P-S in an oversimplified way whereas, if *RPs* is adopted, P-S is reconstructed *as it is seen by the actors*.

Building on Latsis´ distinction between *RPo* and *RPs*, Nadeau (1993) discusses the role of *RP* in Popper (1985) and maintains that *RPo* is clearly false because actors´ behaviour is not always adequate to the `objective´ P-S whereas *RPs* is irrefutable and, hence, it can only be interpreted as a metaphysical statement (*op. cit*., p. 459). He then states that *RPs* is the correct interpretation since `the rationality principle that Popper puts at the theoretical core of all social sciences looks *more* like a "synthetic a priori truth or pure reason" in the domain of social reality than like an empirical law of nature´. Nevertheless, when asked to clarify whether *RP* is a `methodological principle´ or an `empirical conjecture´, Popper explains that `[t]his second case is precisely the one that corresponds to my own view of the status of the rationality principle: I regard the principle of adequacy of action (that is the rationality principle) as an integral part of every, or nearly every, testable social theory´ (Popper, 1994, p. 177). In other words, he views *RP* as an integral part of any empirical theory in the social sciences and, more specifically, as the animating part, just as the laws of motion of planets are an integral part of Newton´s gravitational theory. In an attempt to make sense of all this, Lagueux

[4] In his analysis of the role and status of *RP* Latsis (1983) focuses on Popper´s analysis of the `mind-body problem´, that is, the analysis of the manner in which mental states affect behaviour as discussed in his paper `Of Clouds and Clocks´ (Popper, 1966). According to Latsis (1983, p. 139), *RP* represents Popper´s compromise solution to this problem whereby it is suggested that `our mental states control some of our behaviour and that this control is "of a plastic kind"´

(2006, p. 203) argues that a methodological principle cannot be a part of a scientific theory whose constituent parts must be *empirical* rather than *a priori*. Yet, according to him, if *RP* cannot be a methodological principle, `the *decision* to immunize it can nonetheless be considered as based on a methodological principle´ (see Notturno, 1998, pp. 405-408). Although he uses a different terminology — methodological `rule´ instead of methodological `principle´ — de Bruin (2006, p. 213) also explains that the decision to adopt *RP* is a methodological rule according to which `one should always try to explain human behaviour in terms of reasons´. However, he adds that `there are good reasons to doubt whether the kind of principle of rationality that Popper discusses is really empirical at all… one could as well phrase it as a *metaphysical principle* that all actions have reasons´ (de Bruin, 2006, p. 216, emphasis added).

Despite the fact that some commentators have noted that there are two different notions of rationality in Popper´s philosophy, there has been very little discussion about the relation of *RP* and Popper´s evolutionary theory of knowledge and learning (PTKL). The essence of Popper´s theory of knowledge is that all knowledge is conjectural and that we can *never* prove that a hypothesis is true albeit sometimes it is possible to prove that it is wrong. Likewise, the essence of Popper´s evolutionary theory of learning is that all living organisms (including human beings) learn by virtue of an imperfect and unending process that consists of subjecting their conjectures or hypotheses to trial and eliminating those ones which turn out to be erroneous while keeping *provisionally* those ones that are not falsified (Popper, 1963, p. 312).[5] In other words, our knowledge grows in a `negative´ sense by discarding erroneous conjectures through a process of trial and error-elimination. As a result of it, an implication of PTKL is that the most important feature of knowledge is its fallibility. A second element of Popper´s theory of learning is that the learning process is always *imperfect* insofar as it never reaches an optimum adaptation to the surrounding environment.

Now, a number of critics have referred to the apparent *duality* in Popper´s notion of rationality. To be sure, Popper (1985, p. 365; 1994, p. 181) distinguishes between rationality as a personal attitude — which he defines as the attitude of readiness to correct one´s beliefs when they turn out to be wrong — and his *RP* which, according to him, *has nothing to do* with the assumption that men are rational in this sense. Further,

[5] Interestingly, as mentioned in Langlois (1995, footnote 8), Hayek (1967, p. 84) associates the notion of rationality in economics to the ability to `learn from experience´ which is very close to PTKL.

when he presents *SA* as his methodological proposal for the social sciences he writes that `when we speak of "rational behaviour" or of "irrational behaviour" then we mean behaviour which is, or is not, in accordance with the logic of the situation´ (Popper, [1943*a*] 1966, p. 97; 1944-45, sections 31 & 32). Be that as it may, Kerstenetzky (2009, p. 202) argues that the demarcation line between rationality and irrationality in Popper is the *incorrigibility* of one´s beliefs. That is, human behaviour is `rational´ if we are willing to correct our wrong beliefs and `irrational´ if otherwise. Similarly, Lagueux (2006, p. 202) points out that it is the `tendency to correct oneself by criticism´ that represents true rationality in Popper whereas Vanberg (2002, p. 19) remarks that, on the one hand, Popper presents *RP* as the methodological foundation for the social sciences but, in other parts of his work, he sketches out a framework for the analysis of human behaviour that relies on a different approach at purposeful behaviour that he defines as `conjecture-based problem-solving behaviour´. Finally, the conflict between the view of human agency depicted in PTKL and *SA* is also noted by Oakley (1999, p. 25; 2002, p. 468).

The main purpose of this essay is to analyse the compatibility or otherwise of these two seemingly different notions of rationality in Popper´s work. In the process, we will make five claims. Our first claim is that there is a certain tension between PTKL and *SA* when their relation is analysed from the standpoint of the `rationality of the agents´ whose behaviour the theoretician seeks to capture in a situational model albeit the tension disappears when the relation is treated from the standpoint of the `rationality of the theoretician´. Our second claim is that the nature of the tension between PTKL and *SA* depends on whether the theoretician adopts the `objectivist´ or the `subjectivist´ version of *SA*. In particular, we will argue below that the tension between PTKL and the `subjectivist´ version stems from the fact that, in the latter, it is implicitly assumed that agents´ view of P-S is, at least partially, wrong which implies that agents do not tend to eliminate their mistakes as PTKL posits. By contrast, we will argue that the tension between PTKL and the `objectivist´ version of *SA* stems from the fact that: (i) if agents behave according to PTKL it is not necessarily the case that their decisions will be adequate or appropriate to the `logic of the situation´ because the former *only* implies that agents tend to eliminate their (past) mistakes and, hence, in the wake of changes in the surrounding environment agents´ decisions may be inadequate to the `logic of the (new) situation´, and (ii) adoption of the `objectivist´ *SA* implies *de facto* the imposition of the theoretician´s view of P-S on agents´ but it is unlikely that if agents behave

according to PTKL their view of P-S will converge to the theoreticians´. Our third claim builds on the ideas of Hayek (1943) about the nature of the `facts´ of the social sciences and is that, in the way it is presented by Popper and his commentators, the `objectivist´ version of *SA* represents a *limit* case which presupposes that P-S is (fully) *independent* of agents´ beliefs. Our fourth claim is closely related to the previous one and consists of the idea that, if Hayek´s ideas are accepted, it follows that the natural strategy for social scientists is to seek to reconstruct P-S *as agents´ see it*. Our fifth and last claim is that, unlike what Popper and some of his commentators suggest, the difference between the `objectivist´ and the `subjectivist´ version of *SA* is not that in the former the theoretician reconstructs P-S *as it actually is* whereas in the latter she does it as *agents see it* but, rather that in the former the theoretician seeks to reconstruct P-S *as she sees it* whereas in the latter she does it *as she believes agents see it*. The content of the essay is as follows. The following section introduces PTKL. In section 3, we expound Popper´s *SA*. In section 4, we discuss the duality in Popper´s notion of rationality by addressing: (i) the implications of adopting either the `objectivist´ or the `subjectivist´ version of *SA*, (ii) Hayek´s ideas about the `facts´ of the social sciences, (iii) a reformulation of the `objectivist´ and `subjectivist´ version of *SA* that takes on board Hayek´s ideas on the methodology of the social sciences, and (iv) the division line between `rational´ and `irrational´ behaviour in PTKL and *SA*. Finally, section 5 summarizes and concludes.

2. Popper's evolutionary theory of knowledge and learning

Inductive inference is reasoning from the past observed behaviour of objects to their future behaviour. The 'problem of induction' was originally raised by David Hume (1748) who pondered whether inductive evidence can go beyond the available evidence in order to predict future events. He argued that past evidence can tell us only about past experience. Hume's main argument was that we cannot rationally justify the claim that nature will continue to be uniform merely because it has been in the past, as this is using the sort of reasoning (i.e., induction) that is under question, i.e., it would be *circular reasoning*. Hume (*op. cit*.) also noticed that we *tend* to believe that phenomena behave in a regular fashion, that is, that certain patterns in the behaviour of objects persist into the future.

Next, Popper defines the philosophical 'problem of induction' as the problem of providing a rational justification for the belief that the future will be (largely) like the past (Popper, 1972, p. 2). He identifies two problems in Hume´s criticism of induction:

(i) a logical problem (*HL*), and (ii) a psychological problem (*HP*). First, Hume´s *HL* is whether we are justified in reasoning from repeated instances of which we have some experience to other instances of which we have no experience. Hume´s answer to *HL* is negative no matter how many repetitions of the instances there are. Second, Hume´s *HP* is why, notwithstanding it, all reasonable people believe that instances of which they have no experience at all will conform to those of which they have experience. Hume´s answer to *HP* is that 'the psychological mechanism of association forces them to believe, by custom or habit, that what happened in the past will happen in the future' (*op. cit*., p. 90). According to Popper, this explains why Hume abandoned rationalism and viewed man as a product of blind habit. By contrast, Popper argues that there is no such thing as induction by repetition either in psychology or science:

'We do not act upon repetition or "habit", but upon the best tested of our theories which… are the ones for which we have good rational reasons; not of course good reasons for believing them to be true, but for believing them to be the best available from the point of view of a search for truth or verisimilitude… The central question for Hume was: do we act according to reason or not? And my answer is: Yes.' (*op. cit*., p. 95)

Popper restates Hume´s problem of induction as follows. First, he denies that a theory can be simply justified by assuming the truth of certain observation statements. Rather, he insists that all theories are hypotheses and, hence, they can be overthrown (*op. cit*., p. 13). Further, he states that paradoxically induction is *inductively invalid*, that is, the claim that induction is a legitimate way to acquire (true) knowledge needs to be supported by a 'higher' principle that has, in its turn, been established inductively. But this strategy ultimately leads us into an *infinite regress* insofar as we will endlessly need to resort to a superior principle that has been discovered through induction. Second, he puts forward the proposition that the claim that an explanatory universal theory is false can be justified by the truth of certain observation statements (*op. cit*., p. 7). As the typical example goes, no matter how many white swans we come across, the finding of just one black swan will lead to the rejection of the universal statement 'all swans are white'. Consequently, he urges scientists to construct *severe* tests that help detect false theories so that, by a method of *elimination*, they may eventually hit upon a true theory

even though we can never establish its truth (*op. cit*., pp. 14-15).[6] Thus, he argues that there is an *asymmetry* between verification and falsification; any conjecture may be true or false but even if it turns out to be true, there is no way we can *ever* prove it (*op. cit*., p. 12). According to Popper, the method of science is 'the method of bold conjectures and ingenious and severe attempts to refute them' (*op. cit*., p. 81). Since all theories involve universal statements, we can only 'learn' by proving that our knowledge is false. Specifically, learning takes place either when we reject one´s prior theory or when we are forced to adjust one´s theory in a way that recognizes that in its prior version it was false (*op. cit*, p. 81). In short, Popper´s ideas on scientific methodology can be seen as a sub-product of PTKL:

`Although I shall confine my discussion to the growth of knowledge in science, my remarks are applicable without much change, I believe, to the growth of pre-scientific knowledge also — that is to say, to the general way in which men, and even animals, acquire new factual knowledge about the world. The method of learning by trial and error — of learning from our mistakes — seems to be fundamentally the same whether it is practised by lower or by higher animals, by chimpanzees or by men of science. My interest is not merely in the theory of scientific knowledge, but rather in the theory of knowledge in general. Yet the study of the growth of scientific knowledge is, I believe, the most fruitful way of studying the growth of knowledge in general. For the growth of scientific knowledge may be said to be the growth of ordinary human knowledge *writ large´* (Popper, 1963, p. 216).

and, elsewhere, he writes:

`Organisms evolve by trial and error, and their erroneous trials — their erroneous mutations — are eliminated, as a rule, by the elimination of the organism which is the "carrier" of the error. It is this part of my epistemology that, in man, through the evolution of a descriptive and argumentative language, all this has changed radically. Man has achieved the possibility of being *critical of his own tentative trials, of his own theories*. These theories are no longer incorporated in his organism, or in his

[6] Popper adopted Tarski´s theory that truth is correspondence with the facts or with reality (Popper, 1972, p. 44).

genetic system: they may be formulated in books, or in journals; and they can be critically discussed, and shown to be erroneous, without killing any authors or burning any books: without destroying the "carriers"… *critical reason is the only alternative to violence so far discovered´* (Popper, [1943*b*] 1966, p. 292).

Third, Popper argues that theories are genetically incorporated into all our sense organs and this predisposes us to discriminate *a priori* between relevant or absorbable input and input that can be ignored (Popper, 1972, p.72). For instance, sense organs like the eye *only* react to those selected environmental events which they 'expect'. However, according to him, this prior knowledge cannot be the result of observation; it must be the result of adaptation to the surrounding environment by trial and error. He claims that 99 percent of the knowledge of all living organisms is inborn and incorporated into our biochemical constitution (Popper, 1990, p. 46). Furthermore, he argues that there is no theory-free language to help us interpret external data because primitive theories emerge together with language. Therefore, there is no such thing as pure perception since all languages are *theory-impregnated* (Popper, 1972, p. 145). This leads him to reject any epistemology which chooses our 'direct' observations and perceptions as the starting point; the fact that theories are built into our sense organs implies that 'the epistemology of induction breaks down even before taking its first step' (*op. cit*, p. 146).

Lastly, Popper´s rejection of Hume´s *inductive* theory of beliefs formation leads him to maintain that 'logical' considerations may be duly transferred to 'psychological' considerations. According to him, not only do we reason rationally and thus contrary to the principle of induction, but *we also behave rationally*. He labels this the 'principle of transference' (*op. cit*, p. 6). By applying this conjecture to human psychology he then arrives at the method of *trial and error-elimination* in which the trials correspond to the formation of competing hypotheses whereas the elimination of errors corresponds to the refutation of (false) hypotheses. In other words, he proposes the theory that *individuals do not really think in an inductive way but rather form their beliefs by eliminating false hypotheses*. The theory of knowledge and learning that thus emerges is *evolutionary*. However, such theory implies that adaptation is always 'imperfect':

'Some of the errors that have entered the inheritable constitution of an organism are eliminated by eliminating their bearer; that is, the individual organism. But some

errors escape, and this is one reason why we are all fallible: our adaptation to the environment is never optimal, and it is always imperfect' (Popper, 1990, p. 47).

Further, Popper asserts that *no equilibrium state of adaptation* is reached by the application of the method of trial and error-elimination since (i) no optimal trial solution to any specific problem is likely to be offered, and (ii) the emergence of new structures and instructions involves a continuous *change* in the environmental situation (Popper, 1994, p. 4). More specifically, and crucially, Popper presents the growth of knowledge as bringing in its wake changes in the surrounding environment:

'Our very understanding of the world changes the conditions of the changing world; and so do our wishes, our preferences, our motivations, our hopes, our dreams, our fantasies, our hypotheses, our theories. Even our erroneous theories change the world, although our correct theories may, as a rule, have a more lasting influence. All this amounts to the fact that *determinism is simply mistaken*' (Popper, 1990, p. 17).

In short, Popper makes it clear that the past *affects* but does not determine the future, i.e., the future is not pre-determined. That is, the future is 'objectively open' (*op. cit*., pp. 17-18). As noted in Vanberg (2002, p. 8), Popper´s theory of knowledge and learning exhibits a remarkable similarity to the arguments developed by biologist Mayr (1988) in the sense that both Popper and Mayr argue that intentional problem-solving behaviour can be interpreted as behaviour governed by programs or conjectures which are the product of evolutionary learning by trial and error-elimination. As Vanberg (*op. cit*.) explains, this approach implies that 'there is a continuum from the behaviour of primitive organisms, governed entirely by genetically coded programs, to the sophisticated, deliberated actions of "rational man" governed by conjectures or mental models that are stored in memory'. According to Vanberg (*op. cit*., p. 27), 'even the most deliberate and conscious instances of problem-solving are no less "program-based" than any subconscious or unconscious routine behaviour, in the sense that they, too, have nothing else to rely on than *conjectures*…´. Thus, Vanberg views rationality as a problem-solving capacity that is stored on a person's catalogue of conjectures or programs that exhibits no more wisdom than that embedded in the knowledge acquired in the past. According to this view, rationality 'cannot guarantee pre-adaptedness, it is instead a matter of the backward-looking adaptedness of behavioral programs that

allows for a tentative, forward-looking response to current problem-situations' (*op. cit*., pp. 28-29). As we will see, this aspect of the notion of human rationality embedded in both PTKL and Mayr´s notion of problem-solving behaviour implies the ability to solve certain problems that agents have encountered in the past does not necessarily imply that they are endowed with the ability to solve new (and different) problems that they encounter as the surrounding environment changes.

Let us address Popper´s distinction between *subjective* and *objective* knowledge. The former consists of certain inborn dispositions of organisms and of their acquired modifications to act, whereas the latter consists of the logical content of our theories and, as such, it includes the world of language, conjectures, arguments, and scientific theories.[7] As for subjective knowledge, Popper´s diagnosis is that it is part of a complex but accurate apparatus of adjustment that, in the main, works like objective conjectural knowledge, namely, by the method of trial and error-elimination or 'auto-correction' (Popper, 1972, p. 77). As for objective knowledge he notes that only a tiny part of it can be given sufficient reasons for certain truth. Such tiny part is denoted as *demonstrable* knowledge and comprises the propositions of formal logic, and arithmetic. All else, including knowledge associated to the natural and the social sciences, is conjectural or hypothetical knowledge and, hence, there are no sufficient reasons for holding it to be true (*op. cit*., p. 75). Thus, from the point of view of objective knowledge, all theories are *conjectural* albeit this does not preclude the possibility that some of them are true.

The method of science, according to Popper, is 'the method of bold conjectures and ingenious and severe attempts to refute them' (*op. cit*., p. 81). Since all theories involve universal statements, we can `learn´ by proving that our knowledge is false. Specifically, learning takes place either when we reject one´s prior theory or when we are forced to adjust one´s theory in a way that recognizes that in its prior version it was false (Popper, 1972, p. 81). Thus, we can only 'learn' by refuting our prior knowledge claim. As noted in Boland (2003*b*, p. 242), an implication of PTKL is that the mere accumulation of information does not increase the odds that our theories happen to be true because, as Popper insists, all we can 'learn' from experience is that some of our theories are false. In this respect, Boland (2003*b*, p. 248) makes a useful distinction between the quantitative and qualitative views of knowledge. The former corresponds to

[7] Popper´s notion of `subjective knowledge´ also bears a strong resemblance to the notion of 'reasoning instincts' used in evolutionary psychology (Cosmides & Tooby, 1994, p. 330).

the so-called 'bucket theory of knowledge' whereas the latter corresponds to Popper´s theory of knowledge. He then proposes the metaphor that, in Popper´s Socratic theory of learning, 'knowledge is more like health that one can improve than wealth that one can have more of'. Hence, according to Boland (*op. cit.*), learning consists of improving one´s knowledge rather than of increasing it.

The distinction between objective and subjective knowledge also leads Popper to distinguish between three different worlds or ontological domains: (i) the world of physical objects or states (World 1), (ii) the world of states of consciousness, or of mental states (World 2) and, lastly, (iii) the *autonomous* Platonic-like world of objective contents of thought, especially of scientific thoughts and works of art (World 3).[8] His main thesis on this respect is that almost all our subjective knowledge (belonging to World 2) depends upon World 3, that is, on *linguistically formulated* theories (Popper, 1972, p. 74). However, he argues that our mind may be connected to objects of both World 1 and 3 and this allows World 2 to act as a *mediator* between them. Further, he notes that World 3 exerts a profound influence upon World 1 through the actions of technologists who implement changes in World 1 by applying the predictions of these theories. Finally, he argues that we always select our P-S against a World 3 background which consists of, at least, a language and that 'the activity of understanding consists essentially in operating with third-world objects' (*op. cit*., p. 164). In particular, the development of science and art presupposes the prior existence of the human language which leads Popper (1990) to argue that the latter is, by far, the most important product of the human mind:

'Language makes it possible to consider our theories critically: to look at them as if they were external objects, as if they belonged to the world outside of ourselves which we share with others. Theories become objects of criticism, like the beaver dam' (Popper, 1990, p. 51).

Next, Popper sees science as one of the greatest creations of the human mind, comparable only to the emergence of a descriptive and argumentative language, since

[8] According to Gattei (2009, p. 58), Popper´s World 3 bears a strong similarity to Plato´s theory of Ideas, and to Hegel´s theory of the Objective Spirit although he thinks it is `closer to Bolzano´s theory of a universe of statements in themselves and truth in themselves, or to Frege´s universe of objective contents of thought´.

its creation allowed men to replace: (i) the elimination of error in the violent struggle for life by non-violent rational criticism, and (ii) killing (World 1) and intimidation (World 2) by the impersonal arguments of World 3 (*op. cit*., p. 84). He defines epistemology as the theory of the growth of *scientific knowledge*, that is, the theory of problem-solving, or of the critical discussion, evaluation, and critical testing of competing theories (*op. cit*., p. 142). However, as we mentioned above, PTKL is not only applicable to scientific knowledge but to any type of knowledge. As such, he sees scientists acting on the basis of hunches and guesses concerning what looks promising for future growth in the third world of objective knowledge. In so doing, he identifies content and virtual explanatory power as the most important criteria for the *a priori* appraisal of theories where both are related to their degree of testability. In turn, the most important criterion for their *a posteriori* appraisal is 'verisimilitude' or 'nearness to truth' and this, he argues, depends upon the way a theory has stood up to severe tests (*op. cit*., p. 143).[9] The evaluation process is always critical and aims at error-elimination. Lastly, the exposure of scientific hypotheses to severe tests and criticism from the scientific community guarantees their increasing accuracy at explaining phenomena:

`What is characteristic of science is that the selective system which weeds out among the variety of conjectures involves deliberate contact with the environment through experiment and quantified prediction, designed so that outcomes quite independent of the preferences of the investigator are possible. It is pre-eminently this feature that gives science its greater objectivity and its claim to a cumulative increase in the accuracy with which it describes the world´ (Campbell, 1974, p. 434).

Finally, the growth of World 3 is not a repetitive or cumulative process alike Lamarckian instruction but a Darwinian selection process which consists of systematic error-elimination (*op. cit*., p. 149; also Popper, 1994, ch.1). He identifies three different levels of adaptation: genetic, behavioural learning, and scientific discovery. Scientific discovery is, according to him, a special case of adaptive behavioural learning. Popper asserts that, on all levels, the mechanism of adaptation to the surrounding environment is essentially the same, i.e., a Darwinian selection process by trial and error-elimination. In short, he views science 'as a means used by the human species to adapt itself to the

[9] As he aptly notes, a tautology, though obviously true, has zero *truth content* and zero verisimilitude.

surrounding environment: to invade new environmental niches, and even to invent new environmental niches' (Popper, 1994, p. 2).

3. Popper's methodological prescription for the social sciences

We now address Popper´s methodological prescription for the social sciences known as *SA* and the status of *RP*. Early presentations of the method of *SA* can be found in Popper´s *Open Society* (Popper, [1943*a*]1966, ch. 14, especially p. 97), in his *Poverty of Historicism*, originally published in three articles in *Economica* and, then, as a book (Popper, 1944-45, sections 31 & 32), in a French paper (Popper, 1967), and in *Objective Knowledge* (Popper, 1972, p. 179). However, the place where he presents it thoroughly is in the article titled "Models, Instruments, and Truth: The Status of the Rationality Principle in the Social Sciences" (Popper, 1994, ch. 8). This book chapter was originally written in response to an invitation that Popper received in the early 1960s from the Department of Economics at Harvard University and the lecture he delivered there on 26 February 1963. As noted in de Bruin (2006, footnote 1), in 1963 and 1964 two new sections were added and a small extract was then circulated in the London School of Economics in 1967 and 1968. This extract was translated into French and published as `La rationalitè et le statut du principe de rationalitè´ (Popper, 1967) and, then, a Spanish translation of the French translation appeared about a year later. A revised version of the English extract was published in 1983 on pages 357-365 of an anthology titled *A Pocket Popper* which is currently available in *Popper Selections* (Popper, 1985). However, the full text of the speech at Harvard University was not made available until 1994 when it was published in a collection of Popper´s essays titled *The Myth of the Framework*.

3.1. The Rationality Principle

Popper´s thesis in that chapter is that there is no fundamental difference between the natural sciences and the social sciences since both of them resort to the construction of models or *typical* P-S to explain and predict events. If anything, models are viewed by him as being even more important in the social sciences due to the non-existence of universal laws. In any case, he argues that the models of the theoretical social sciences are always an *over-simplification* of reality and, hence, do not represent the facts truly. According to him, the fundamental problem of the social sciences is 'to explain and understand events in terms of human actions and social situations' (Popper, 1994, p. 166). In turn, the reconstruction of social situations should include the consideration of

the relevant 'social institutions' which he defines as 'all those things which set limits or create obstacles to our movements and actions' (*op. cit*., p. 167). In his autobiography, Popper makes it clear that his methodological proposal for the social sciences stems from an 'attempt to generalize the method of economic theory (marginal utility theory) so as to become applicable to the other theoretical social sciences' (Popper 1976*a*, pp. 117-118).

Next, Popper makes a distinction between 'rationality' as a personal attitude and his *RP*. In particular, he makes it clear that his *RP* has nothing to do with the assumption that men adopt a rational attitude. Rather, he defines it as an *a priori* methodological principle which assumes that *our actions are adequate to our problem-situations as we see them* (Popper, 1994, p. 181). More specifically, he remarks that *RP* is *not* true: 'The rationality principle is false. I think there is no way out of this. Consequently, I must deny that it is *a priori* valid' (Popper, 1985, p. 361).[10] Notwithstanding it, he believes it represents a good approximation to the truth. Thus, *RP* 'does not play the role of an empirical explanatory theory, of a testable hypothesis' (*op. cit*., p. 360). Rather, he views it as an integral part of every testable theory and proposes to avoid blaming it whenever our theory breaks down in the wake of empirical tests. His methodological advice to social scientists is thus never to abandon *RP* so that, in the wake of a refutation of their model, they should always revise their models of the agent's P-S.[11]

As Koertge (1975) shows, Popper's views on the *RP* have evolved over time. As time passed by, he tended to weaken his claims about the kinds of actions that agents could be expected to perform so that 'where he had earlier spoken of actions as being `rational´ or `appropriate´, he now characterized them as `adequate´, or `adapted´, or `in accordance with the situation´ (*op. cit*. p. 441). According to him, the most likely reason for this evolution in terminology was his increasing emphasis on the fact that the P-S which played a central role in the explanation was not so much the agent's *objective* P-S but, rather, the agent's *theory* of her P-S or the P-S *as the agent saw it* (Koertge, 1975,

[10] Of course, this applies to any premise aimed at providing a 'closure' for a model. As Loasby (1999, p. 14) reminds us, 'all closures are in some degree false. There can be no self-sufficient Cartesian scheme for deducing justified true knowledge from some original certainty'.

[11] This touches upon the issue of the incompatibility of Popper´s *RP* and his falsificationist methodology. For instance, Caldwell (1991, p. 13) argues that 'Popper´s rationality principle represents an immunizing stratagem that is elevated to the status of an inviolable methodological principle'. In an attempt to reconcile both principles, Koertge (1979, p. 93) interprets *RP* as the Lakatosian hard-core of Popper´s research program in the social sciences whereas the positive heuristic is 'his metaphysical theory of man as an evolving rational problem-solving animal'.

p. 442). She explains that *RP* really consists of two clauses: the first (*RP-1*) says that 'every action (by a person) is a rational response to some problem-situation' whereas the second (*RP-2*) tells us that 'every person in a problem-situation responds rationally to it' (*op. cit*., p. 443). In turn, *RP-1* entails: (i) that the response was issued through a *methodical* appraisal of the set of possible solutions available to the actor, (ii) that a description of both P-S and the appraisal process could be verbalized by the actor, and (iii) that the person acted as she did as a result of the appraisal process so that if a better alternative had been available to her she would have taken it. Thus, the complete *RP* formulated in Koertge (1975) emphasizes the close connection between the action and the *systematic deliberation* process from initial conditions that made the agent behave as she did. Further, Koertge (1979, p. 90) points out that requirement (i) above implies that 'some systematic non-random decision procedure be used' albeit she notices that Popper did not specify the minimal requirements which acceptable decision rules should satisfy. This means that *RP* can, in principle, be supplemented with different theories of belief formation. As Koertge (*op. cit*., p. 92) explains, for Popper, to explain an action using *RP* does not 'imply that the agent's beliefs are reasonable nor even that her way of making decisions is the best possible one' but only presupposes that agents assess the situation in a *systematic* way.

Next, as we noted above, Popper´s methodological advice to social scientists is never to abandon *RP* so that, in the wake of a refutation of their model, they should always revise instead their model of the agent's P-S. He offers two arguments in favour of this strategy: (i) that we learn more if we blame our situational model, and (ii) that the adoption of *RP* 'reduces considerably the arbitrariness of our models' (*op. cit*., p. 362). As for the first argument, he explains that:

'The main argument in favour of this policy is that our model is far more interesting and informative, and far better testable, than the principle of the adequacy of our actions. We do not learn much in learning that this is not strictly true: we know this already' (Popper, 1985, p. 362).

Likewise, Caldwell (1991, p. 25) argues that, although immunizing stratagems should be generally avoided 'at least in the special case of situational analyses, one is able to *criticize more severely and obtain fruitful criticisms* if one blames the model

rather than *RP* whenever a falsification occurs'. As for the second argument, Popper explains that:

'The attempt to replace the rationality principle by another one seems to lead to complete arbitrariness in our model-building. And we must not forget that we can test a theory only as a whole, and that the test consists in finding the better of two competing theories which may have much in common; and most of them have the rationality principle in common' (Popper, 1985, p. 362).

As Hands (1985, p. 87) remarks, Popper's first argument above means that if we are consistent with *RP* 'the falsification of a specific theory only means that we have misspecified the "situation", i.e., that we have attributed the wrong preferences or constraints to the individual'. In turn, Popper's second argument implies that although *RP* is potentially falsifiable *we choose to make a methodological decision that, when faced with a falsifying observation, we will stick to it* and revise instead our hypotheses about the desires, beliefs, and constraints faced by agents (Hands, 1985, p. 88). Notably, Becker (1976) resorts to a similar argument to justify the use of rational choice theory. According to him, human behaviour can be viewed from the standpoint of individuals who seek to maximize their utility from a stable set of preference and subject to a given constraint. Where action appears to deviate from the predictions of neoclassical utility theory, Becker claims that little is gained from resorting to explanations in terms of irrationality, changes in preferences or cultural values, etc… for such explanations are *ad hoc* and may even be contradictory. Furthermore, he adds that the question is left unanswered of just *why* human behaviour should be sometimes rational but sometimes not.

According to Caldwell (1991, p. 15), there are two main weaknesses in Popper´s presentation of *SA*: (i) vagueness about how it should be implemented, and (ii) Popper´s apparent belief that *SA* is the only adequate method to adopt in the theoretical social sciences. As for the first point, we have presented above a clearer explanation of how to apply it suggested in Koertge (1979). As for the second point, Caldwell (1991, p. 16) readily admits that *SA* is a powerful and fruitful method for the social sciences, yet he criticizes Popper´s idea that *SA* is the *only* legitimate method for the theoretical social sciences. Further, he recognizes that there is a tension between falsificationism and *SA* owing to the fact that *RP* adopts the status of a methodological prescription that plays

the role of an immunizing stratagem.[12] A discussion of this issue is in Hands (1985, p. 89) who argues that, by Popperian standards, scientific explanations based on *RP* 'are as close to metaphysical explanations as they are to scientific explanations' and, hence, the tension between these two methodological principles can hardly be resolved. Caldwell (1985) proposes to solve the conflict between *SA* and falsificationism by adopting a broader conception of acceptable scientific practice based on 'critical rationalism' whose goal is to subject theories to an *optimal* amount of criticism. In turn, the latter will depend on both the specific problem to be solved and the nature of the problem under investigation (*op. cit*., p. 25). Such prescription was proposed in Klappholz & Agassi (1959) and, later on, it has been promoted by Boland (2003*a*) who stresses that the only generally applicable methodological rule is the exhortation to be always critical and ready to subject one's hypotheses to critical scrutiny. More specifically, he insists we should focus on the Socratic-Popper identified in Klappholz & Agassi (1959) and thus discard the Lakatos-Popper (also known as Popper the 'falsificationist') promoted by Latsis (1972) and Blaug (1975). According to him, if we put falsificationism aside in favour of 'critical rationalism' the conflict between *SA* and falsificationism vanishes.

3.2. The two versions of the Rationality Principle

Latsis (1983) was probably the first commentator to identify the existence of an `objectivist´ version (*RPo*) and a `subjectivist´ version (*RPs*) of *RP* in Popper´s work.[13] In the former, the relevant P-S is that one as seen by the theoretician whereas, in the latter, the theoretician is supposed to reconstruct P-S as seen by agents. Latsis (*op. cit*.) denotes the former as the `strong´ version of *RP*. According to Latsis (*op. cit*., p. 131), Popper both *weakens* and *widens* the notion of rationality in human behaviour when adopting *RPs*.[14] Building on the distinction between *RPo* and *RPs*, Nadeau (1993, p. 463) notes that `an attentive reading of the 1967 text shows that although Popper views the *RP* as an explanatory principle throughout the text, he surreptitiously changes his

[12] Notwithstanding the intellectual authority of Popper, economists like Hayek (1967, p. 29) or Hutchison (1977, p. 43) question the applicability of Popper's falsificationanist methodology to economics. Indeed, the applicability of strict falsificationism to the social sciences appears to be problematic even to Popper himself as noted in Hands (1985, p. 96).

[13] However, we will argue below that a clear antecedent of this distinction is in Hayek (1937).

[14] By contrast, we will argue below that the `subjectivist´ version of *RP* is a legitimate and potentially fruitful one in the social sciences.

way of formulating it during the course of his argument, going from an objectivist formulation at the beginning of the text to a subjectivist formulation at the end´. Hands (1991, footnote 14) recognizes that `Popper is really unclear on this´, and Latsis (1983, p. 133) claims that Popper seems either `confused or deliberately elusive´ on this issue. Be that as it may, Hands (*op. cit*.) points out that in his 1985 text Popper adopts the subjectivist interpretation when he openly says that rationality is only `as agents see it´ and *SA* can thus be applied to apparently irrational behaviour such as the behaviour of a `madman´ (Popper, 1985, p. 363).[15] However, he adds that Popper also denotes *SA* `a purely objective method´ which `can be developed independently of all subjective and psychological ideas (Popper, 1976, p. 172) and that, elsewhere, Popper says that *RP* is the `general law that *sane* persons as a rule act more or less rationally´ (Popper, 1966, p. 265).

Now, in a passage of his 1967 French paper, Popper (1985, p. 363) proposes his famous example of the `flustered driver´ who, by trying to park stubbornly his car in evidently too small a space, does not act in a way that is appropriate to the situation *in which he finds himself* and then recognizes that `we employ the rationality principle to the limit of what is possible whenever we try to understand the action of a madman´ (Popper, 1994, p. 179). It is in the section of the chapter where he notices that cases of neurosis have been explained by Freud and other psychologists with the help of their own version of the *RP* that he switches to a *subjectivist* version of *RP*. Then, in a key note to one of the paragraphs (footnote 19), he acknowledges that he refers successively to two versions of his *RP* and even identifies a third intermediate version according to which P-S is said to be `as the agent could (within the objective situation) have seen it´ (Popper, 1994, ch. 8, footnote 19).[16] In the aftermath of it, Lagueux (2006, p. 201) concludes that, according to Popper, `what the agent sees may or may not be considered a part of the objective situation that the model describes´. Summing up, the `objectivist´ version (*RPo*) supposes that agents possess `true´ knowledge; the `subjectivist´ version

[15] Yet, Zouboulakis (2014, p. 87) argues that it is clear that Popper has an `objectivist´ version of *RP* in mind.

[16] Prior to this clarification, Popper (1972) had already recognized the existence of two versions of *SA*:

> `There are many cases in which we can reconstruct, *objectively* (even though conjecturally), (a) the *situation as it was* and (b) a very different *situation as it appeared to the agent*, or as it was *understood, or interpreted* by the agent. It is interesting that this can be dome even in the history of science´ (Popper, 1972, p. 179, footnote 27).

(*RPs*) supposes that the alleged knowledge that agents possess is *partially* wrong; and the third version constitutes an intermediate case. However, in all three versions of *RP* it is assumed that the agent acts in a way that is appropriate to the state of *his* knowledge (Popper, 1994, ch. 8, footnote 19; Lagueux, 2006, p. 201).

Next, building on the terminology coined in Latsis (1972), Kerstenetzky (2009, p. 201) denotes *RPo* the `maximal´ or `single-exit´ interpretation and *RPs* the `minimal´ or `multiple-exit´ interpretation of *RP*. The `single-exit´ interpretation stems from the fact that, if it is supposed that the agent perceives P-S in an objective way, there is thus only `one´ possible solution whereas the `multiple-exit´ interpretation captures the idea that, in principle, there are as many solutions as subjective perceptions of P-S exist. It is the `multiple-exit´ interpretation that is of interest in the context of the `subjectivist´ *SA*. In particular, the issue is *whether* we can assume for methodological purposes that the different subjective perceptions of P-S held by actors actually converge on a `single´ one and, if so, *how* this convergence comes about. Alike Jacobs (1990), Kerstenetzky (*op. cit*) associates the `objectivist´ or `single-exit´ modelling to the influence on Popper of the work of Weber. By contrast, Hedström *et al*. (1998, p. 359) do not think there is textual evidence that Popper got the inspiration for the notion of *SA* from Weber´s work and suggest that if there was any influence at all it was probably *indirect* since Hayek — a friend of Popper — admired Weber. However, it could be argued that Popper´s method for the theoretical social sciences takes on board Weber´s notion of `interpretive understanding´ or `*verstehen´* — developed later on by the Austrian economists — and, especially, his notions of `ideal type´ and of `instrumental rationality´, i.e., the use of rationality to bring about change in the surrounding world in the interest of the actor (Weber, 1949).[17] Be that as it may, there is some textual evidence that points to Hayek as the most important *direct* source of influence on Popper´s work. Notably, Popper (1994, ch. 8, note 1) writes that `I was particularly impressed by Hayek´s formulation that economics is the "logic of choice"´ as expressed in his essay titled `Economics and

[17] However, it has been argued that, although the notion of `single-exit´ modelling and *SA* originates with Weber, for him this and related concepts were tools of *historical* analysis, not of theory (Langlois, 1995, p. 230; 1998, p. 69). Interestingly, a philosophical foundation for Weber´s notion of `ideal type´ is in the school of phenomenology and, particularly, in Schutz´s concept of `second-order typifications´ (Schutz, 1972).

Knowledge´ (Hayek, 1937, pp. 33ff). According to Popper, it was this that led him to the formulation of the `logic of the situation´ in his *Poverty of Historicism*.[18]

Finally, there is the issue of the status of *RP*. We have already mentioned above the profound ambiguity of Popper´s explanation about the status of RP. The subsequent discussion about the role and status of *RP* among Popper´s commentators focused on the distinction between *RPo* and *RPs*. For instance, Lagueux (1993, 2006) argues that, even if we adopt the `subjectivist´ interpretation, *RP* cannot be *a priori* true because, according to him, it is simply not true that people always act appropriately according to the P-S *as they see it*. Notwithstanding it, he thinks that *RP* occupies an exceptional place in the social sciences because it constitutes a *condition of intelligibility* of any phenomenon that derives from human action. More specifically, the latter can only be intelligible, i.e., understood by an external observer, when it is motivated by reasons, that is, when it represents an appropriate response to P-S as seen by the agent (Lagueux, 2006, p. 205). He concludes that maintaining *RP* after acknowledging that it is not *a priori* true is, after all, to claim that `in spite of the fact that irrational decisions occur, human actions are nonetheless normally understandable´ (*op. cit.*).

4. The notion of rationality in Popper´s philosophy of the social sciences

In section 2 we showed that PTKL implies that: (i) all knowledge is conjectural, (ii) that we learn through an (endless) process whereby we subject our conjectures to trial and discard those ones that turn out to be wrong, and (iii) that the learning process is imperfect and never converges to an optimum. Consequently, the most important feature of knowledge is its fallibility. By contrast, Popper´s methodological proposal for the social sciences has been denoted as 'situational determinism' (Latsis, 1972; Oakley, 2002) which suggests that there may be some key epistemological differences between PTKL and *SA*. The first thing we should like to note is that PTKL is a theory about the nature of knowledge and its growth over time while *SA* is a methodological prescription

[18] The interrelation between the ideas of Hayek and Popper is discussed in Oakley (1999) and Caldwell (2003). For instance, Popper (1943) appears to have been inspired by Hayek (1942) when expressing his crucial idea that both our institutions and traditions are largely the `*indirect, the unintended and often the unwanted by-products*´ of conscious and intentional human actions and, therefore, that `only a minority of social institutions are consciously designed, while the vast majority have just "grown", as the undesigned results of human actions´ (Popper, [1943*a*] 1966, p. 93). Likewise, Hayek had already argued that social studies deal `not with the relations between things, but with the relations between men and things or the relations between man and man. They are concerned with man´s actions and their aim is to explain the unintended or undesigned results of the actions of many men´ (Hayek, 1942, p. 276).

aimed, arguably, at speeding up the rate of progress of the social sciences so that these two elements of Popper´s philosophy do correspond to the positive and methodological domain respectively. That said, we believe there is also a normative element in PTKL since trial and error-elimination can also be said to be the way we *should* behave when seeking to expand our knowledge. In any case, error-elimination can only proceed *after* there is clear-cut evidence that, retrospectively, a decision made in the past was wrong. However, this does not provide us with a systematic rule for making decisions *in the future* other than to avoid repeating the same mistakes made in the past. In short, error-elimination is an incomplete guide to decision-making.

Next, we may wonder how *SA* would look like if the agents that are the object of the modelling exercise exhibited a theory of knowledge and learning akin to PTKL. To be sure, the situational model of the typical P-S consists of three elements: (i) external (and observable) elements such as the physical and social constraints agents are subject to, (ii) the knowledge and information that agents possess, and (iii) their goals and aims. Now, if agents behave according to PTKL, then the situational model of the typical P-S should incorporate the knowledge they possess which would include the experience accumulated from mistakes they made in the past given the specific circumstances that prevailed at that time. Therefore, adequate behaviour would imply, as a minimum, not repeating previous mistakes. However, as we noted above, there is no further guidance for agents stemming from PTKL as far as future decision-making is concerned in case they encounter new (and different) P-S. In short, PTKL appears to be compatible with *SA* provided the situational model includes agents´ learning from previous mistakes.

4.1. PTKL versus *SA*: the `rationality of agents´

We noted above that several commentators, as well as Popper (1994, ch. 8, note 19) himself, identify two different versions of *RP*: an `objectivist´ version (*RPo*) and a `subjectivist´ version (*RPs*). According to the former, the relevant P-S is the `objective´ P-S, that is, the P-S *as it actually is* whereas, according to the latter, the theoretician should reconstruct P-S *as it is actually seen by the agents*. As Popper (1972, p. 179) recognizes, in both cases P-S is *conjectured*.[19] That said, we will argue below that, if Hayek´s ideas on the nature of the `facts of the social sciences´ are taken on board, there

[19] In this respect, let us mention that Menger ([1871] 1950, p. 148) was one of the first social scientists to incorporate error within his model and to argue that *all* knowledge (both of the actors and of the theorist) is bound to be prone to errors.

is no reason *a priori* to expect that the theoretician´s view of P-S is *closer* to the `true´ P-S than agents´ (Hayek, 1943). This is because, as Hayek argues, the theoretician does not possess superior relevant knowledge that is not shared by agents. Be that as it may, *RPo* and *RPs* constitute two different modelling strategies in the social sciences the consequences of which, to the best of our knowledge, have not been explored so far. Notably, an antecedent is Schumpeter´s distinction between `objective rationality´ and `subjective rationality´ (Schumpeter, 1984). He defines the former as consisting of the `applicability of a rational schema to the actor´s behaviour´ and he defines the latter as the `conformity of the actors´ mental processes to a rational schema´ (*op. cit*., p. 583). Crucially, he states that the former need not imply the latter and criticizes the tendency of some social scientists to *implicitly* identify the rationality of the `observer´ with the subjective rationality of the `observed´ (*op. cit*., p. 583). He uses the example of the neoclassical theory of monopoly to illustrate the notion of `objective rationality´:

`The model just described is the product of the analyst´s mind as much as any physical theory is, and does not in itself say anything about reality or about anybody´s actual behavior or rationality… Even if the model should fit anyone´s behaviour this does not mean that the individual in question consciously aims at the result and still less that he arrives at it by processes at all similar to the analytic procedure´ (Schumpeter, *op. cit*., p. 580).

Schumpeter´s notion of `objective rationality´ is closely associated to his notion of `rationality of the observer´ whereas the notion of `subjective rationality´ is coupled to his notion of the `rationality in the observed´. In the example of monopoly theory, he explains that the construction of a model will give us the conditions under which the maximization of profits will be attained thereby setting up a standard against which the theorist can compare actual behaviour. However, he makes it clear that such model is *entirely* a product of the `rationality of the observer´ and, hence, the usefulness of the modelling exercise will depend on the degree to which that hypothesis is justified by facts (*op. cit*., p. 580). According to him, a common source of divergence between the type of human behaviour that stems from the `rationality of the observer´ and the `rationality in the observed´ is the existence of a multiplicity of ends in actors´ minds. To the extent that the goals of actors are also an element of P-S, the adoption of a subjectivist interpretation of *RP* will require that the theoretician *understands* the goals

of actors without this necessarily implying that she shares them. Now, the relevance of his notion of `subjective rationality´ emerges clearly in those cases where the situational model constructed on the basis of the `rationality of the observer´ *does not fit the facts*. As Schumpeter notes, in such cases the task of the theoretician is to explain the reasons for the *discrepancy* between the `rationality of the observer´ and the `rationality in the observed´ (*op. cit*., p. 586). In turn, this will require an effort by the former to adopt the point of view of the `observed´:

`Understanding an end and judging rationality of means often requires that the analyst "puts himself" into places very far distant from his time and social location. Sometimes he has to transplant himself into another cultural world´ (Schumpeter, 1984, p. 583).

These ideas on the methodology of the social sciences were originally written by Schumpeter *circa* 1940 for a Harvard discussion group on rationality which included Parsons, Leontief, and Sweezy. The manuscript remained unpublished for more than 40 years until Professor Loring Allen of the University of Missouri in St. Louis found it among the papers of Schumpeter in the Harvard University archives. It was published posthumously in 1984 at the *Journal of Institutional and Theoretical Economics*. We believe this manuscript contains some intuitions that exhibit a high degree of affinity with Popper´s notions of *RPo* and *RPs* (Popper, 1994). However, we should like to note that it was Hayek´s *Economica* essay `Economics & Knowledge´ (Hayek, 1937) where the distinction between `objective´ and `subjective´ rationality was first formulated. In that essay, Hayek criticises equilibrium economic theory for making an illegitimate use of the concept of `data´ possessed by economic agents as well as for the methodological confusion thus created:

`But this does not solve the question whether the facts referred to are supposed to be given to the observing economist, or to the persons whose actions he wants to explain, and if to the latter, whether it is assumed that the same facts are known to all the different persons in the system, or whether the "data" for the different persons may be different… There seems to be no possible doubt that these two concepts of "data", on the one hand in the sense of the objective real facts, as the observing economist is supposed to know them, and on the other hand in the subjective sense, as things known

to the persons whose behaviour we try to explain are really *fundamentally different and ought to be kept carefully apart*. And, as we shall see, the question why the data in the subjective sense of the term should ever come to correspond to the objective data is one of the main problems we have to answer´ (Hayek, 1937, p. 39, emphasis added).

We know that Popper had read Hayek´s 1937 paper in *Economica* and, indeed, he refers to it as the key source of his understanding of the core of economics (Popper, 1994, p. 181, footnote 1). According to Popper, it was Hayek´s exposition of the `logic of choice´ in that paper that led him to the formulation of the `logic of the situation´ as embracing both the `logic of choice´ and the `logic of historical P-S´. Yet, Popper does not refer explicitly to Hayek´s distinction between subjective and objective data. That said, it is very likely that Popper´s recognition, later on, of a distinction between *RPo* and *RPs* is related to his acquaintance with Hayek´s *Economica* essay. In the following sections we will explore in some detail the relation between these two concepts as well as their relation to PTKL *from the point of view of the agents* that are the object of the modelling exercise performed by the theoretician.

4.1.1. PTKL *versus* the `subjectivist´ version of *SA*

Let us focus on the relation between *RPs* and PTKL. To be sure, *RPs* constitutes a *minimal* requirement for rationality in that agents´ behaviour only has to be adequate or appropriate to the P-S *as they see it*. This implies that, as in the Austrian School of von Mises, Hayek, and Schumpeter, rationality is associated to behaviour that is *goal-directed* or *purposive*.[20] This type of rationality is sometimes denoted as *instrumental* in the sense that reason becomes an instrument to reach a certain goal, e.g. an increase in pleasure.[21] Although there are significant methodological differences among members of the Austrian School of economics, they all viewed economics as part of a science of

[20] As von Mises explains:

`Every human action aims at the substitution of more satisfactory conditions for less satisfactory. Man acts because he feels uneasy and believes that he has the power to relieve to some extent his uneasinesss by influencing the course of events. A man perfectly content with the state of his affairs would not have any incentive to change things; he would have neither wishes nor desires, he would not act because he would be perfectly happy... Strictly speaking, only the increase in satisfaction (decrease of uneasiness) should be called *end*, and accordingly all states which bring about such an increase *means´* (von Mises, 1944, p. 532).

[21] The notion of `instrumental rationality´ goes back as far as Max Weber´s sociological histories of world religions culminating with his classic study of modern European Christianity (Weber, 1904-5).

human action whose core is `to be found in the unique property possessed by human beings of engaging in operations designed to attain a state of affairs that is preferred to that which has hitherto prevailed´ (Kirzner, 1976, p. 148). What is crucial in our context is that the Austrian School´s conception of rationality is *subjective* in the sense of being an *a priori* assumption about human behaviour. There are two sources of subjectivity. First, there is the subjectivity of actors´ ends or, as von Mises puts it:

`Nobody else than the individual himself can decide what satisfies him better and what less... There is no such thing as an absolute state of satisfaction or happiness irrespective of the desires of the individual concerned´ (von Mises, 1944, p. 533).

Second, there is the subjectivity of knowledge itself in the social sciences. As long argued in Hayek (1943), it is only in the social sciences that our *interpretation* of a situation no matter whether it is right or wrong becomes an integral part of the situation thus affecting subsequent developments. Further, and to the extent that we understand the surrounding world via the 'internal models' we create, our understanding of the world will affect our decisions and, in this way, it may affect the world itself. Hayek (*op. cit.*) illustrates this theme by explaining the purposive nature of human action. As he explains, just as we cannot speak of the objective properties of a tool without saying something about the purpose for which the tool is used so we cannot speak of social institutions objectively. Laws and economic institutions cannot be known apart from the intentions of the individuals who use them. In the field of economics, for instance, the value of money depends on the opinions of individuals who use it rather than on any inherent property of it. As we argue below, Hayek´s ideas on the methodology of the social sciences seem to have been ignored by most commentators of Popper´s work in that field.

Members of the Austrian School of economics like von Mises or Hayek adopted the `praxeological approach´ which consists of a theory of human action based on a set of self-evidently true *a priori* axioms on behaviour which, in turn, yields conclusions which are true regardless of time and place. However, the axioms of praxeology are not arbitrary like, for instance, those of mathematics. Rather, Austrian economists maintain that these axioms are already given to us in our minds and that, through the exercise of `introspection´ or `verstehen´, which consists of understanding the functioning of our minds, we have the possibility of understanding the behaviour of others. That said, the

extreme subjectivism of the Austrian School of Economics leads to the conclusion that there is no possibility of acquiring knowledge about any social phenomena other than through `introspection´. Further, the notion of rationality proposed by von Mises (1944) as *purposive* behaviour may preclude the generation of predictions which can be subject to empirical tests. This is because the hypotheses about social phenomena derived from self-evident axioms may be close to being true but they may also possess little empirical content. This problem is addressed in Popper (1963, pp. 217-19) who makes it clear that science characterises as preferable `the theory which tells us more; that is to say, the theory which contains the greater amount of empirical information or content´. In other words, the empirical content of a theory increases with the increasing *improbability* of it being true or else with its increasing exposure to falsification. He uses the example of meteorological forecasts; a forecast according to which in some unspecified time in the future it will rain has, as he explains, a high probability of being true yet it has virtually no empirical content, whereas a forecast which specifies the date and the time it is likely to rain has a high degree of empirical content yet it is quite likely to be false. Likewise, predictions derived from general or self-evident axioms on human behaviour like the ones of praxeology — which amount to stating little more than all human behaviour is purposeful — have a high probability of being true yet they have little empirical content because they are not falsifiable. By contrast, the subjectivist version of *SA* proposed by Popper (1985) is not subject to the previous criticism since, in addition to incorporating all the relevant elements of P-S — including the physical and social constrains and the knowledge and information possessed by agents — it also posits that actors´ behaviour is `adequate´ to P-S *as they see it*. The requirement that actors´ behaviour is `adequate´ — in addition to being purposeful or goal-oriented — implies, in turn, that the empirical content of theories constructed upon *RPs exceeds* the empirical content of theories about human behaviour derived from praxeology.

Let us distinguish between 'means-rationality', 'beliefs-rationality', and 'ends-rationality' (Hamlin, 1986). 'Means-rationality' implies the *correctness* of one´s actions *given* one´s desires and beliefs regardless of whether the latter are right. Therefore, as a minimum 'means-rationality' implies *consistency* of choice by agents. For instance, in neoclassical microeconomics, 'means-rationality' is characterized by consistency in the preferences of households or *transitivity*: if an agent prefers *a* to *b* and *b* to *c*, then *a*

must also be preferred to *c*.[22] The further requirements that are usually imposed, i.e., that individuals´ preferences exhibit both 'completeness' and 'continuity', are not ones of 'means-rationality' but rather of the optimization methods through which economists seek to represent individual preferences by a 'utility function'. Thus, when economists speak of `rational´ agents what they usually have in mind is that their choices have to be, at least, consistent with one another.

Next, following Hamlin (*op. cit*), 'beliefs-rationality' implies that an individual´s (subjective) model of the surrounding world represents a good enough approximation to reality. Similarly, Bicchieri (1992) defines 'epistemic' rationality as a characteristic of beliefs that consists in their being *correct* given the evidence that is available to agents. Finally, 'ends-rationality' means that the behaviour of agents is *purposeful* or oriented to the achievement of a goal and, hence, not the result of chance (Hamlin, *op. cit*.). For instance, in neoclassical economics 'ends-rationality' is associated to the pursuit of *self-interest*. To be sure, this has been so since Edgeworth who, in his 1881 *Mathematical Psychics*, stated that 'the first principle of Economics is that every agent is actuated only by self-interest'. Prior to the emergence of 'neoclassical' economics in the second half of the 19th century, the classical statement of mainstream economics methodology is in Mill ([1836] 1967). Be that as it may, Twomey (1998, p. 435) claims that the clearest statements of this tradition were already present in the works of Bentham and Hobbes. In particular, he argues that Bentham (1789) first formulated the notion that agents are motivated to maximise pleasure whereas Hobbes (1651) provided a statement of egoism according to which individuals *always* seek their own greatest good. Therefore, we may characterise the `praxeologic´ models of the Austrian school of economics as implying `ends-rationality´ and the situational models based upon *RPs* as implying both 'means-rationality' and 'ends-rationality' but *not* 'beliefs-rationality'. As we will argue below, each of these types of models implies a different division line between `rational´ and `irrational´ behaviour.

The absence of 'beliefs-rationality' in models based on *RPs* implies that agents´ beliefs may well be wrong *ex-post*, i.e., that agents are fallible. Specifically, agents may

[22] For instance, in the neoclassical theory of the consumer, 'means-rationality' consists of the axioms of completeness, independence, reflexivity, and transitivity of preferences (Dow, 1995, p. 724). In the case of the expected utility model developed by von Neumman and Morgenstern (1947), which constitutes the modern theory of choice under risk, 'means-rationality' is satisfied if the following four assumptions or axioms are fulfilled: cancellation, transitivity, dominance, and invariance. To this, we may add the more technical assumptions of comparability and continuity.

perceive the physical and social constraints they face erroneously or may simply possess wrong information. Therefore, the adoption of *RPs* implies that agents´ beliefs are liable to error and, hence, understanding their behaviour (including their mistakes) will require the construction of a model of the typical P-S *as seen by agents*. It follows from this that the adoption of *RPs* is *a priori* compatible with PTKL since the agents in the typical model can make mistakes stemming from their wrong beliefs. Yet, the notion of adequate behaviour according to PTKL also implies, as noted above, the requirement that agents `learn´ from their past mistakes, i.e., they do not repeat them. Consequently, full compatibility of *RPs* with PTKL would require that the theoretician recognizes that agents do not repeat their mistakes in the future. As we will see below, this feature of PTKL may create a tension with *RPs* when the purpose of the modelling exercise is to make predictions.[23] We may also add that, if *RPs* is adopted, the point of view of the theoretician *vis-à-vis* the actors becomes analogous to the position of participants in the `Beauty Contests´ that were so popular in the British tabloids in the 1930s and that were metaphorically captured by Keynes in his *General Theory* to explain the formation of prices in financial markets (Keynes, 1936, p. 156). To be sure, in `Beauty Contests´, what participants were supposed to do in order to win the prize was not so much to identify — among the photos of beautiful ladies portrayed in a panel — the lady they believed to be the most beautiful one but to `guess´ the photo of the lady they believed other participants would select as the most beautiful one. In a similar fashion, we will argue below that, if *RPs* is adopted, the theoretician seeks to reconstruct P-S not as she sees it herself but *the way she thinks agents see it.*

Next, and crucially, to the extent a *discrepancy* exists between the theoretician´s (objective) view of P-S and his conjecture about agents´ view of P-S, the generation of predictions will require making the crucial assumption that *such a discrepancy and the*

[23] As noted in Beinhocker (2013), in the context of the social sciences, a prediction (unlike a forecast) amounts barely to the *deductive logical consequences* of a theory. It should be noted, however, that the predictions to be derived from the models in the social sciences differ from the predictions generated by the theories of the natural sciences. In particular, Popper (1994, p. 163) distinguishes between explaining or predicting *singular events* from problems of explaining or predicting a *kind* or *type* of event. According to Popper, the former can be solved *without constructing a model* — no more than certain universal laws and the relevant initial conditions are needed — whereas the latter is most easily solved *by means of constructing a model* (*op. cit*., p. 164). Further, he argues (Popper, *op. cit*., p. 165), that a model consists of `certain elements placed in a typical relationship to each other, plus certain universal laws of interaction — the "animating" laws´. Unlike theories, models try to capture the *typical* aspects of P-S in order to make statements about a *type* of event and, hence, they represent something akin to *typical initial conditions* (*op. cit*., p. 164). In turn, a statement about a *typical* event can be either an explanation of *why* that typical event occurred in the past, or else, a prediction, that is, a logical consequence of the theory.

*`situational factors´ that warrant it exhibit a high degree of stability over time*. In turn, the former implies that the `null hypothesis´ in empirical tests applied on a situational model which adopts *RPs* is that agents´ view of P-S is, at least partially, *wrong* whereas the alternative hypothesis is that the theoretician´s (objective) view of P-S is correct. If such discrepancy were to disappear over time for some reason (e.g., learning by agents), the predictions derived from it would be equivalent to the predictions generated if *RPo* were adopted. Thus, we disagree with Vanberg (2002, p. 12) when he argues that the `subjectivist´ *RP* poses a testability problem vis-à-vis the `objectivist´ *RP*. Specifically, situational models that adopt *RPs* can generate predictions albeit, as we argued above, their generation implicitly implies adopting the assumption that agents´ view of P-S *will remain constant in the future*. Unless the theoretician does so, the models´ testability will be seriously weakened. This is because, in the wake of an unfavourable empirical test, the theoretician may try to circumvent its refutation by arguing that the adverse result of the empirical test was due, for instance, to an (unpredictable) *change* in agents´ view of P-S. Thus, a *sine qua non* condition for potential refutability in this case, i.e., for the model to possess `empirical content´, is that agents´ view of P-S is assumed to remain constant over time or, else, that agents follow a constant pattern of behaviour in spite of the mistakes such behaviour may bring about. However, this assumption creates some tension with PTKL since, according to the latter, agents tend to purge their wrong beliefs over time.

Now, we have argued above that, if *RPs* is adopted, it is implicitly assumed that (i) there is a discrepancy between the theoretician´s view of P-S and agents´ view of P-S, and (ii) that the former persists over time. As we argued above, this implies (under the null hypothesis) that the agents whose behaviour the theoretician seeks to capture in the situational model *do not revise their wrong beliefs* which runs counter to PTKL. The discrepancy alluded to above is between two different conjectures: (i) the theoretician´s (objective) view of P-S, and (ii) agents´ view of P-S. According to Popper (1994, p. 178), the latter is always part of the former since the theoretician can only understand agents´ view of P-S if she reconstructs a *wider* view of P-S than their own. Specifically, if we adopt Popper´s interpretation of *RPs*, what the theoretician subjects to empirical test is the hypothesis that agents *systematically* fail to perceive the `true´ P-S and, under the null hypothesis, this implies that agents´ view of P-S will be disappointed *if* the theoretician´s view of P-S is correct. However, as we have argued above, the systematic disappointment of beliefs will come about because it is implicitly assumed that agents

do not `learn´ from their mistakes, i.e., they tend to repeat mistakes all over. However, the above-mentioned tension between PTKL and *RPs* does not arise if the main purpose of constructing a situational model is to *explain the past* (e.g., historical interpretation) rather than to generate predictions. This is because in the former case the theoretician need not be concerned about the persistence into the future of a discrepancy between agents´ view of P-S and her `objective´ view of P-S. In short, *RPs* is more problematic than Popper recognized if the purpose of constructing a situational model is to generate predictions. [24]

To finish off this section, a clear example of this tension between PTKL and *RPs* is the Keynesian-type business cycle theory proposed by Minsky (1975) which is based on overoptimistic expectations of economic agents about their ability to honour future cash commitments which result from their inherited liability structure. According to Minsky´s `financial instability hypothesis´ (*op. cit*.), market economies are intrinsically unstable owing to the fact that economic agents become *systematically* overoptimistic during the upswing which makes them take on an excessive amount of debt and this eventually triggers off an asset price deflation and a subsequent financial crisis that precipitates the economy into a downswing. In other words, Minsky´s theory posits that, as memories from the last financial crisis fade out, agents will tend to underestimate the risk implied by the increase in the level of real indebtedness so that the upswing ends up when an external factor, e.g., an increase in interest rates, leads to an initial decrease in the price of real and financial assets which then brings about a reassessment of liability structures and, finally, leads to an asset price deflation. In other words, Minsky´s theory is an example of business cycle theory where (i): there is a discrepancy between the theoretician´s `objective´ view of P-S and agents´ view of P-S, and (ii) it is (implicitly)

---

[24] Notturno (1998, p. 412, emphasis added) argues that the `problem of situational analysis in the theoretical and historical social sciences, in Popper´s view, is not to construct models that predict or prophesize the future; it is to construct models that help us to *explain and understand the past´*. In this respect, Popper seems to view *RPs* as being particularly suited to the task of historical explanation:

`The historian´s task is, therefore, so to reconstruct the problem situation as it appeared to the agent, that the actions of the agent become *adequate* to the situation… Our conjectural reconstruction of the situation may be a real historical discovery. It may explain any aspect of history so far unexplained´ (Popper, 1972, p. 189).

Furthermore, Popper (*op. cit*., p. 166, emphasis in original) explains that `the fundamental problem of both the theoretical and the historical social sciences is *to explain and understand events in terms of human actions and* [typical] *social situations´* and, hence, does not mention the generation of *predictions* as one of the aims of the social sciences.

assumed that agents tend to repeat their past mistakes so the same phenomenon (i.e., business cycles), occurs recurrently and inevitably.

4.1.2. PTKL *versus* the `objectivist´ version of *SA*

According to Hands (1992, p. 28), `it is easy to see that situational analysis is the method of microeconomics (and of any macroeconomics based on micro foundations)´. Indeed, Popper recognizes that his source of inspiration for *SA* is the methodology of neoclassical microeconomics (Popper, [1943*a*] 1966, p. 97; 1944-45, p. 82; 1976*a*, p. 93; 1976*b*, 117f).[25] However, on those few occasions when Popper makes it clear that he intends to extend the methodology of neoclassical economics to the rest of the social sciences he seems to have in mind the `objectivist´ *SA*. That said, some commentators have noticed that the rationality requirements are more demanding if *RPo* rather than *RPs* is adopted (Latsis, 1983; Farmer, 1998; Oakley, 1999; Vanberg, 2002). Notably, and in addition to fulfilment of 'means-rationality' and 'ends-rationality', *RPo* implies the fulfilment of 'beliefs-rationality'. The combination of these three types of rationality yields a type of rationality known as 'substantive rationality' (*SR*) (Simon, 1976). *SR* is a notion of rationality that is concerned solely with the *consequences* or outcomes of rational choice. Specifically, Simon (1976, p.130) denotes human behaviour as being substantively rational 'when it is appropriate to the achievement of given goals within the limits imposed by given conditions and constraints'. Thus, as with classical decision theory, the interest lies not so much in *how* decisions are made but in *what* decisions are made. In short, *SR* constitutes a special case of *RPo* in so far as, in addition to consistent and purposeful behaviour, it is assumed that *agents´ beliefs are correct on average*. As we have noted above, it is this approach to human rationality that lies at the core of the maximization assumption in neoclassical economics. Indeed, some scholars have argued that *RPo* is the key principle that underlies the methodology of mainstream economics (Farmer, 1998; Oakley, 1999).[26]

---

[25] It is clear that Popper admired Neoclassical Economics. For instance, he writes that:

`The social sciences never had for me the same attraction as the theoretical natural sciences. In fact, the only theoretical social science which appealed to me was economics´ (Popper, 1976*a*, p. 121).

Notwithstanding it, Blaug (1985, p. 287) argues that `Popper knew little about social sciences and less about economics´.

[26] For instance, Farmer (1998, p. 27) argues that *RP* is in the `hard core´ of the (Lakatosian) economist´s research programme. Matzner & Jarvie (1998) even suggest that Popper´s *SA* represents a *soft* version of `economic imperialism´ to be distinguished from a *strong* version they associate with the work of Gary

4.1.2.1. Rationalizing the `objectivist´ version of *SA*

Now, one can rationalize *RPo* as a *methodological decision* according to which the theoretician assumes beforehand that the mistakes made by agents (by `mistakes´ we mean decisions that are adequate from the point of view of P-S as seen by the agents but *inadequate* from the viewpoint of the P-S as seen by the theoretician) *are declared to be less interesting for the purpose of understanding agents´ behaviour and, especially, for generating predictions than the modelling mistakes made by the theoretician.* In other words, a rationale for *RPo* is that, although agents´ mistakes cannot be ruled out *a priori* — so *RPo* would be compatible with fallibility — nevertheless the theoretician *chooses to ignore the former for methodological reasons*. What are these reasons? First, that the theoretician *gains little*, if anything, by learning that agents make mistakes because (i) she already knows it and, more importantly, (ii) that the nature of the mistakes agents make is likely to change over time in an *unpredictable* way and so learning about them is of little help for the purpose of generating predictions. To be sure, learning about the mistakes agents have made in the past may be helpful if we have the assurance that the *same* mistakes (i.e., mistakes triggered off by the very same factors) will be repeated in the future. As we explained above, it is this scenario that may justify the adoption of *RPs*. However, if such condition is not satisfied, there is arguably little we can learn from the mistakes made by agents in the past other than for the purpose of historical analysis. A second reason for adopting this methodological decision would be that, if that decision were not adopted then, in the wake of an erroneous prediction of the model the theoretician might be tempted to *sidestep* its falsification by arguing that agents´ beliefs and decisions were, on a particular episode — the one that was subject to the test — different from what one would `objectively´ expect and to utilize this `anomaly´ as a justification for the (adverse) result of the empirical test. By contrast, if *RPo* is adopted, the *onus of proof* will inescapably rest on the theoretician´s view of P-S.

According to the rationale for *RPo* we have suggested above, ascribing a role to agents´ errors in the situational model (i.e., to discrepancies between their view of P-S and the theoreticians´ view of P-S) will prevent us from generating predictions unless agents´ errors are predictable. Yet, one possible reason why agents´ errors may actually be unpredictable is that agents may `learn´ from their past mistakes so that their future

---

Becker. In any case, they recognize that it was Popper — and not Gary Becker — who first formulated a programme for extending the logic of economics to the noneconomic social sciences (*op. cit*., p. 336).

mistakes will tend to *differ* from previous ones. Thus, in order to generate predictions, a hypothesis which assumes that agent´s view of P-S does not coincide with the theorist´s view of it (i.e., *RPs*) will need to be coupled to an additional assumption according to which agents´ errors tend to persist over time and, hence, are *predictable*. However, to the extent that this assumption implies that agents do not `learn´ from their mistakes, it is in conflict with PTKL. To be sure, if agents `learn´ from their previous mistakes so that they do not repeat them, their future mistakes will tend to *differ* from their previous mistakes and, unless the range of potential mistakes is limited, their future mistakes will thus be unpredictable. By contrast, if the theoretician adopts *RPo* instead of *RPs*, this problem does not arise because there is no presumption that agents´ view of P-S differs from the theoretician´s. As we will argue below, the adoption of *RPo* actually implies the *imposition* of the theoretician´s view of P-S upon agents. Does this mean that there is no tension between the `objectivist´ version of *SA* and PTKL? We believe not. Firstly, PTKL only implies that agents `learn´ by trial and error-elimination so that behaviour that is rational according to PTKL is not necessarily appropriate to the `logic of the situation´ faced by agents. To be sure, the errors that agents made in the past occurred in the environment that surrounded them at that time so if the latter changes in an unpredictable way agents may well make new (and different) errors, i.e., they may make decisions that are not appropriate to the `logic of the situation´. Secondly, since the theoretician´s view of P-S does not necessarily coincide with agents´ view of it, several further assumptions will need to be made to justify the *coincidence* of the theoreticians´ and agents´ views. These additional assumptions are presented and discussed below. In any case, we may anticipate that these assumptions are problematic in the sense that if it is assumed that agents behave according to PTKL it is doubtful that their decisions will be appropriate to the `logic of the situation´ as it is seen by the theoretician, even on average. In short, the nature of tension between PTKL and *RPs* differs from the nature of tension between PTKL and *RPo* in that, in the former case, tension stems from the fact that the adoption of *RPs* implies that agents tend to *repeat* the same mistakes they made in the past in a way that is hardly compatible with `learning´ by trial and error-elimination as posited in PTKL whereas, in the latter case, it is simply assumed that agents´ view of P-S coincides, at least on average, with the theoretician´s view of P-S so that their behaviour is always appropriate to the objective `logic of the situation´ yet, as we will argue below, the mechanisms by virtue of which such coincidence is justified can hardly be reconciled with PTKL.

4.1.2.2. How do the views of the agents and the theoretician tend to converge?

Next, we have suggested above that the adoption of *RPo* implies *de facto* the *imposition* of the theoreticians´ view of P-S upon agents´. However, scientists do not usually see things this way. An example is mainstream economics where a number of mechanisms have been suggested in the literature to (implicitly) justify the adoption of *RPo*. To be sure, such mechanisms are seemingly viewed by mainstream economists as reasons why they need not care about agents´ beliefs when reconstructing P-S because an `objective´ P-S can be said to exist `out there´ that is *sufficiently* independent of agents´ beliefs so that the theoretician can reconstruct P-S as she sees it. There are two mechanisms through which the neglect of agents´ beliefs by the theoretician is normally justified: (i) the operation of the `law of large numbers´ in the social domain, and (ii) the presence of `learning´ by individuals. However, as we show below both mechanisms are problematic. Let us address the first mechanism. According to it, agents´ decisions may turn out to be objectively wrong in retrospect but nevertheless their *mistakes will tend to cancel each other out provided the number of individuals is large enough*. The former implies that the scope for fallibility at the aggregate level in this version of *RPo* is negligible since it is restricted to random mistakes associated to transitory factors. There is textual evidence which suggests that some influential social scientists have implicitly resorted to this mechanism to justify the adoption of *RPo*. For instance, Nobel Laureate in Economics John Hicks (1956, p. 55) writes that `the preference hypothesis [in the context of neoclassical utility theory] only acquires a *prima facie* plausibility when it is applied to a statistical average´. More explicitly, Gibbard & Varian (1978) describe optimizing behaviour by individuals as capturing the 'central tendency' of economic behaviour or:

'If deviations are random or more precisely, are not systematic, there might be good reason to have some faith in the conclusions of the [economic] model even though the assumptions, strictly interpreted, are implausible. Perhaps a case in point is the economist´s assumption of perfect optimizing behaviour. Of course, this assumption is strictly speaking, false, but, so long as errors in optimization are not systematic, this hypothesis may be useful in describing the "central tendency" of economic behaviour. Furthermore, in models where individual units´ behaviour is being aggregated, non-systematic errors may be expected to "wash out" in the process of aggregation' (*op. cit.*, p. 670).

Likewise, it has been argued in the sphere of sociology that it is not necessary to claim that all agents optimize but, instead, that the tendency to optimize is the most important *non-idiosyncratic* factor at work so that the operation of a sort of 'law of large numbers' guarantees that optimizing behaviour dominates (Goldthorpe, 1998, p. 169). We believe this assumption (i.e., the `law of large numbers´) implicitly lies at the core of neoclassical economics where agents are modelled *as if* they were infallible – when they exhibit perfect foresight – or as if their mistakes were random (Muth, 1961). More specifically, this assumption is implied when the optimizing assumption is be applied in modelling exercises. However, the `law of large numbers´ in statistics assumes that the different trials of a stochastic process are: (i) independent and, crucially, (ii) have the *same* distribution so that, as the number of such trials tends to infinity, the probability distribution of a random variable *concentrates* around the finite expected value of each of the trials.

Now, it is unlikely that these conditions will be satisfied in the case of agents´ view of P-S. For one thing, there are likely to be significant interdependencies among agents´ (subjective) view of P-S owing to the presence of conventional elements so that condition (i) is likely to be violated. Further, agents´ (subjective) view of P-S may differ substantially from others agents´ which also violates condition (ii). Thus, the conditions for reliance on the `law of large numbers´ as it exists in statistics for the purpose of providing a rationale for the coincidence, on average, between agents´ view of P-S and the theoreticians´ view of P-S are not warranted. Consequently, the adoption of *RPo* can only be justified on *strict* methodological grounds. That said, we believe that some advocates of *RPo* assume that if the theorist´s view of P-S diverges significantly from agents´ view of P-S such discrepancy will tend to be eliminated over time by means of other mechanisms such as: (i) trial and error-elimination, and (ii) imitation of successful strategies by agents. In other words, advocates of *RPo* may argue that the occurrence of learning at the individual level based on trial and error-elimination or the imitation of the successful strategies of others will make agents´ view of P-S eventually converge to the theorist´s (objective) view of P-S so that, for the sake of analytical convenience, we may freely assume that agents´ beliefs are correct. However, the imitation of successful strategies requires that some other agents have previously `learnt´ to perform some tasks adequately so that the presence of some kind of learning is a *sine qua non* condition for the imitation of successful strategies to allow some other agents to make decisions that are appropriate to the `logic of the situation´. Thus, for the sake of simplicity, we will

leave aside the latter. In other words, the theorist assumes in this case that the operation of a *negative* feedback mechanism whereby agents systematically revise their beliefs until the latter coincide with the theoretician´s view of P-S justifies the adoption of the methodological decision to assume that agents´ beliefs are eventually correct. In short, `learning´ is the second mechanism (additional to the `law of large numbers´) by means of which mainstream economists may try to justify the assumption that agents´ view of P-S coincides with the theorist´s.[27] For instance, Nobel Laureate R. Lucas characterizes the type of situations on which economic theory focuses as the *end-result* of an adaptive learning process:[28]

---

[27] An alternative rationale to agents´ learning process for the adoption of the optimization assumption in mainstream economics is the 'natural' selection argument. This argument was originally proposed by Alchian (1950) in the context of economic competition and since then it has been advocated by a number of neoclassical theorists. Alchian (*op. cit.*) views the economic system as an *adoptive* mechanism which cleverly chooses among actions generated by the adaptive pursuit of `profits´. His argument starts with the premise that the realization of profits is the criterion according to which successful firms are selected. However, he admits that this process may be *independent* of the nature of the decision-making processes of economic agents. In particular, he recognizes that economic success does not require proper motivation but may instead be the outcome of fortuitous circumstances. Further, to the extent that profits accrue only to those firms who are better than their competitors, what matters for survival is one's competitiveness *relative* to the others rather than their absolute proximity to an optimum. Nevertheless, he recognizes that conscious or purposeful behaviour (in addition to sheer chance) also plays a role in the selection of firms by the market. Two such examples of purposeful behaviour are: (i) imitation of strategies previously adopted by successful firms, and (ii) the adoption of trial and error strategies aimed at improving one's adaptation to the environment. However, he makes it clear that the former should *not* be understood as mechanisms through which adequate or `rational´ actions can be selected thereby allowing firms to converge to an optimum in the form of profit maximization. This is because the latter would require the fulfilment of the two following convergence conditions (Alchian, 1950, p. 219): (i) that every single trial must be classifiable either as a success or as a failure, and (ii) the continual rising toward some *optimum optimorum* without the occurrence of intervening descents. In this respect, he writes:

'These convergence conditions do not apply to a changing environment, for there can be no observable comparison of the result of an action with any other… As a consequence, the measure of goodness of actions in anything except a tolerable-intolerable sense is lost, and the possibility of an individual´s converging to the optimum activity via a trial-and-error process disappears. Trial and error becomes survival or death. It cannot serve as a basis of the *individual´s* method of convergence to a "maximum" or optimum position.' (*op. cit.*).

The upshot of Alchian´s discussion is that successful adaptation within a stable environment may give *ex-post* the appearance of rational or optimizing behaviour at the individual level *even though no ex-ante rational calculation actually occurred*. Further, and as Loasby (1999, pp. 20-21; see also Vromen, 1995, pp. 32-33) insists, the economic `survival´ argument can only allow its advocates to claim that surviving firms will have achieved results which are, on average, *closer* to the maximisation of profits than those firms that did not survive and, hence, it does not allow them to use it as a justification for the occurrence of optimization at the *individual* level as Friedman (1953, p. 22), for instance, does. Although with significant qualifications, developments of Alchian´s argument can be found in Friedman (1953) and Becker (1962) and critical assessments of the `survival´ argument are in Vromen (1995), Loasby (1999) and Lagueux (2010).

[28] By contrast, Arrow (1986, p. S385) criticizes the view that the optimizing assumption can be justified as being the result of a process of learning and adaptation.

'Economics has tended to focus on situations in which the agent can be expected to "know" or to have learned the consequences of different actions so that his observed choices reveal stable features of his underlying preferences… Technically, I think of economics as studying decision rules that are steady states of some adaptive process, decision rules that are found to work over a range of situations and hence are no longer revised appreciably as more experience accumulates…' (Lucas, 1986, p. 218).

However, for this feedback mechanism to be *effective*, it is necessary that: (i) the former is fast and accurate enough, and (ii) P-S remains constant until the convergence process has been completed.[29] Yet, as Tversky and Kahneman (1986, p. 90) insist, such conditions rarely arise in the real world. In particular, the former can hardly be satisfied when agents make decisions in a changing environment in which it is hard to ascertain whether an observed outcome is a direct consequence of our decisions or a consequence of someone else´s decisions. Furthermore, Popper (1994, p. 4) insists that *no optimal state of adaptation* is ever reached by the application of the method of trial and error-elimination owing to: (i) the continuous *change* in the environmental situation, and (ii) agents´ inability to eliminate all their errors. That said, we believe that it is the reliance on the alleged efficacy of `learning´ at the individual level that makes some scientists implicitly assume that any negative result that occurs in the wake of empirical tests can be ascribed only (or mainly) to their *own* modelling mistakes (i.e., to their own failure to capture the `objective´ P-S properly) rather than to agents´ mistakes. Thus, although *RPo* accounts for the presence of learning by agents, it exhibits some clear differences with PTKL in that `learning´ at the individual level in the latter is unlikely to warrant that agents´ decisions will be adequate to the `logic of the situation´ as it is *seen by the theoretician*.

4.1.3. The dichotomy between `rational´ and `irrational´ behaviour

The notion of rationality that is concerned with the *consequences* or outcomes of rational choice is known as 'substantive' rationality (SR). Specifically, Simon (1976, p. 130) identifies human behaviour as being substantively rational 'when it is appropriate

[29] Wible (1984-85, p. 271) characterizes the approach to decision-making and expectations formation embedded in neoclassical economics as one of *instantaneous rational assessment* and hypothesizes that its origin is the emphasis in the 'logic of justification' of knowledge rather in the 'process of discovery' of knowledge made by proponents of the logical positivist school of philosophy.

to the achievement of given goals within the limits imposed by given conditions and constraints'. According to this definition, rational behaviour is a type of purposeful or intentional behaviour directed towards a goal, e.g. the maximization of utility. As with classical decision theory, the interest lies not so much in *how* decisions are made but in *what* decisions are made. According to Simon (1965, p. 84), theoretical models based on 'substantively' rational individuals share a common framework characterised by: (i) a set of alternative courses of action that are available to the individual, (ii) (perfect) knowledge that permits the individual to predict the precise consequences of choosing any possible alternative[30], and (iii) a criterion for determining *which* set of potential consequences she prefers. In such models, rationality is usually defined as 'the ability of actors to select that course of action which leads to the most preferred set of predicted consequences' (*op. cit.*). SR assumes that the surrounding environment is either known or knowable (i.e., the stochastic environment is stable), and individuals have sufficient cognitive abilities to deal with a complex reality. SR is the type of rationality actors are assumed to exhibit in those models that adopt strong versions of *RPo* such as the ones that prevail, for instance, in mainstream economics. In particular, agents who exhibit SR must fulfil ends-rationality, means-rationality, and beliefs-rationality. This is shown in the third row of Table 1 below. In turn, this implies that behaviour that falls short of maximizing is deemed 'irrational' (Becker, 1962). Specifically, a violation by agents of either means-rationality or beliefs-rationality is interpreted in mainstream economics as signalling `irrational´ behaviour. Likewise, Popper ([1943*a*] 1966, p. 97) explains that `when we speak of "rational behaviour" or of "irrational behaviour" then we mean behaviour which is, or which is not, in accordance with the logic of the situation´. Thus, Popper´s notion of rationality in the context of *SA* bears, arguably, a strong resemblance to the notion of rationality in mainstream economics.

That said, we believe that the charge of `irrationality´ is a direct implication of the adoption of *RPo* and, particularly, of the *imposition* upon agents of the theoretician´s view of P-S. More specifically, we believe that the `irrationality´ charge that is applied to those agents who fail to maximize (a given pre-specified objective function) obeys ultimately to a failure to distinguish between the `rationality of the theoretician´ and the `rationality of agents´. For instance, the implicit assumption by mainstream economists

[30] In the absence of perfect foresight, the notion of 'substantive' rationality usually requires the existence of a stable stochastic environment that allows individuals to confidently extrapolate the past into the future as in, for instance, 'subjective expected utility' (SEU) theory.

that agents´ subjective view of P-S *coincides*, at least on average, with the theoretician´s view of P-S logically implies that behaviour that falls short of the rationality standard ascribed to the theoretician is `irrational´. However, we believe that if agents´ view of P-S does not coincide with the theoretician´s then the former cannot be blamed for being `irrational´. In particular, an individual cannot be said to be `irrational´ if her beliefs are incorrect. Rather, as PTKL has it, she can only be said to be `irrational´ if she refuses to revise her (wrong) beliefs.

The counterpart to SR is the notion of 'procedural' rationality (PR). According to Simon (1976, p. 131), 'behavior is procedurally rational when it is the outcome of appropriate deliberation'.[31] PR shifts attention from the consequences of choice to the *process* of choice where the emphasis is placed in the presence of a decision process based on the use of simple heuristics or 'rules of thumb'.[32] PR can thus be characterised as the ability of actors to use simple heuristics that are *adequate* for a specific purpose. Reliance on simple heuristics to make decisions assumes that, most of the time, actors face situations characterised by (i) Knightian uncertainty (Knight, [1921]1971), or (ii) where the 'optimal' solution is intractable. The former corresponds to scenarios where either we do not have an exhaustive list of potential consequences of a certain decision or else to situations where, even if such list were available, it is impossible to attach numerical probabilities to them. In turn, the latter corresponds to situations where there are insurmountable constraints on the ability of agents: (i) to identify optimal actions given a set of beliefs and desires, and (ii) to acquire the information that is relevant to the problem at hand. In contrast, SR implies that agents make decisions by following the prescriptions of Bayes's rule or by maximizing expected utility, as in SEU theory.[33]

[31] Note the similarity of this definition with Koertge´s (1975) reformulation of *RP* provided above.

[32] PR is coupled to its sister notion of 'bounded' rationality (BR). According to Simon (1979), economic agents´ knowledge is subject to three different types of constraints: (i) limited ability to process, analyse, and store information, (ii) uncertainty, and (iii) the presence of social institutions. BR stems from the fact that the existence of these constraints prevents economic agents from 'optimizing'. Simon coined the term 'satisficing' to denote a decision-making rule that attempts to meet an acceptability (minimum) threshold. In contrast to 'satisficing' behaviour, the purpose of optimal decision-making is to find the best option available.

[33] According to Volz & Gigerenzer (2012, p. 1), most of the time we make decisions under Knightian uncertainty, while situations of known risk are relatively rare and found mostly in gambling. They note that Savage (1954, p. 16), one of the fathers of the theory of choice under uncertainty, made it clear that applying Bayesian theory to decisions in uncertain worlds would not make sense because there is no way to know all alternatives, consequences, and probabilities. Likewise, Arrow (2004, p. 54) argues that in uncertain, ill-specified worlds, maximization of expected utility 'has no meaning at all'.

Now, we believe PR captures properly the type of rationality implied by PTKL. As we noted above, the acquisition of knowledge in Popper's account runs parallel to the *process* of adaptation to a partly unknown (and changing) environment in which some of the errors made in the past by individuals are purged by virtue of a learning process that consists essentially of subjecting their hypotheses or conjectures to trial and error-elimination. Crucially, Popper emphasises that our adaptation to the surrounding environment is often successful and often unsuccessful. More specifically, some errors will *escape* and this possibility is one of the reasons our knowledge is always fallible (Popper, 1990, p. 47). In turn, Popper (1994, p. 4) insists that the systematic application of the method of trial and error-elimination will not result in an 'optimum' adaptation to the surrounding environment. Instead, and due to the partial elimination of errors, our adaptation to the environment is always *imperfect*. It follows from this that the observed states of adaptation *can never be the result of convergence to an optimum*. Should *all* errors be systematically purged and the environment be stable, the process of adaptation to the latter would eventually be perfect and only then could the state of adaptation be interpreted as the outcome of a convergence to an optimum. In this scenario, individuals would be fallible only to the extent that the changes in the environment cannot be fully anticipated. The method of trial and error-elimination thus consists of an adaptation mechanism whereby errors tend to be eliminated and new hypotheses are then subject to trial. As some commentators note (Kerstenetzky, 2009; Lagueux, 2006), the watershed between `rational´ and `irrational´ behaviour in PTKL is marked by the *unwillingness of agents to correct their wrong beliefs* or, as it were, by the *incorrigibility* of their beliefs. This is clearly stated by Popper in the following quotation:

`The main distinction, I suggest, is that a healthy person´s beliefs are not incorrigible: a healthy person shows a certain readiness to correct his beliefs. He may do so only reluctantly, yet he is nevertheless ready to correct his views under the pressure of events, of the opinions held by others, and of critical arguments… the mentality of the man with definitely fixed views, the "committed" man, is akin to that of the madman… but *in so far as he is committed, he is not rational´* (Popper, 1985, p. 364; 1994, p. 180, emphasis added).

This suggests, as noted above, that there are two different notions of rationality in Popper´s work: (i) behaviour that is in accordance with the `logic of the situation´

(Popper, [1943*a*] 1966, p. 97; 1944-45, sections 31 & 32), and (ii) willingness to revise one´s wrong beliefs (Popper, 1985, p. 364. We have argued above that there is a certain tension between these two notions of rationality when the relation is approached from the standpoint of the `rationality of the agents´ and we will argue below that there is no such tension when the relation is approached from the standpoint of the `rationality of the theoretician´.

Next, unlike models based on *RPo*, both praxeology and models based on *RPs* imply that agents´ view of P-S may be mistaken and, hence, that their decisions may turn out to be wrong *ex-post*. In the case of *RPs*, the theoretician is assumed to adopt the viewpoint of actors and, thus, she is supposed to be able to distinguish between her own view of P-S (i.e., the `rationality of the observer´) and agents´ view of P-S (i.e., the `rationality in the observed´). Further, in the case of models based on *RPs*, agents are assumed to exhibit means-rationality (in addition to `ends-rationality´) which implies that their mistakes can only be ascribed to wrong beliefs and not to an inconsistent or inadequate behaviour *given* the information available to them. However, as far as agents are concerned, there is no mechanism that ensures an adequate, let alone an efficient, use by them of the available information regardless of the accuracy of the latter. Rather, decision-making in a context of pervasive uncertainty can only be the result of a process of systematic deliberation by agents. In other words, although we cannot rule out that a *given* correct decision is the outcome of sheer chance, it is much more likely that correct decisions given the knowledge and information that agents possess will be the result of agents´ systematic deliberation. Therefore, the type of rationality agents exhibit if *RPs* is adopted can be denoted as `procedural´. In turn, this implies that PTKL and *RPs* share the feature that agents are *fallible*, i.e., their beliefs may be wrong *ex-post*. By contrast, as we explained above, agents´ fallibility is negligible in case *RPo* is adopted since, *a priori*, individuals can make mistakes but their mistakes are assumed to cancel out at the aggregate level.

Now, by adopting *RPs* the theoretician seeks to identify agents´ partially wrong beliefs and thus to explain their behaviour accordingly (i.e., by stressing the divergence of agents´ behaviour from what one would expect if their beliefs were correct). By contrast, it is quite unlikely that the theorist can provide a `rational´ reconstruction of agents´ apparently wrong behaviour by pointing instead to their *inadequate* behaviour given their *correct* beliefs (i.e., to absence of means-rationality). To be sure, to provide such an account of agents´ inadequate behaviour by appealing to the notion that some

agents exhibit, for instance, miopic behaviour or weak (or lack of) will, implies a large element of psychologism and, hence, a loss of inter-subjectivity, accountability, and transparency in theoretical analysis. Further, such a diagnose of inadequate behaviour as based on a violation of the assumption of means-rationality may apply to some or even to very few individuals but certainly it will not apply to the majority of them which, according to us, precludes its use in the social sciences. This suggests that, even though it is *not* true that agents always make an adequate use of the information that is available to them and, to the extent that we accept as universally valid the assumption that agents´ behaviour is always goal-oriented, it is unclear whether we can speak of a dichotomy between `rational´ and `irrational´ behaviour by agents in models which adopt *RPs*. By contrast, to the extent that agents only exhibit ends-rationality in praxeologic models, we cannot speak of the existence of such a dichotomy in the latter. All this is illustrated in Table 1 below.

Approach	**Beliefs**	**Rationality**	**Dichotomy**
Praxeology	Right or wrong *ex-post*	Ends-rationality (instrumental)	No
Subjectivist SA	Right or wrong *ex-post*	Ends & means-rationality (procedural)	?
Objectivist SA	Right `on average´	Ends, means & beliefs-rationality (substantive)	Irrationality $\Rightarrow$ failure of either means or beliefs-rationality (or both)
PTKL	Right or wrong *ex-post*	Ends-rationality (procedural)	Irrationality $\Rightarrow$ incorrigibility

Table 1. Classification of approaches to rationality

Finally, let us note that it is the existence or otherwise of a dichotomy between `rational´ and `irrational´ behaviour by agents in the context of *SA* that was the object of an exchange between Nadeau (1993) and Lagueux (1993, 2010). In their contributions, which focus on the epistemological status of *RP*, Nadeau (1993) first argued that *RPs* is the correct interpretation of Popper´s Rationality Principle and subsequently argued that the former is a metaphysical statement and, as such, it is *a priori* true and irrefutable. By

contrast, Lagueux (1993) held that *RP* should rather be interpreted as a methodological principle and, hence, that it is false but approximately true in the sense there may be some instances where it does not hold. Nevertheless, he (as Popper does) believes that *RP* is a sufficiently good approximation to the truth and defends a sort of `statistical´ justification of *RP* in so far as he points out that `the rationality principle can be said to be "approximately true" only to the extent that it applies to a large number of cases´ (Lagueux, *op. cit*., p. 475). Both authors make use of Popper´s example of the `flustered driver´ to substantiate their arguments:

`For the rationality principle seems to me to be clearly false — even in its weakest zero formulation, which may be put like this: "Agents always act in a manner appropriate to the situation in which they find themselves."… In think one can see easily that this is not so. One has only to observe flustered drivers trying to get out of a traffic jam, or desperately trying to park their cars when there is hardly any parking space to be found, or none at all, in order to see that we do not always act in accordance with the rationality principle… Moreover, there are, obviously, vast personal differences, not only in knowledge and skill — these are part of the situation — but also in assessing or understanding a situation; and this means that *some people will act appropriately and others not´* (Popper, 1985, p. 361; 1994, p. 172, emphasis added).

Building on this, Lagueux comments:

`Suppose that we try to explain the [Popper´s] example of the flustered driver using the rationality principle as understood by Popper, that is to say, as a principle that *is empirical and false*. It is clear that, in such a case, the rationality principle *could* be held responsible for the failure of an explanatory theory that was supported by it´ (Lagueux, 1993, p. 475).

By contrast, Nadeau (1993) writes:

`It seems to me to be evident that the *RP*s is logically irrefutable, in exactly the same way that, for Popper, probabilistic assertions or metaphysical statements are irrefutable... My critique of Popper can be clarified further by a brief examination of his example of the flustered driver... Popper uses this example to falsify the *RP*, but how

does this example work exactly? It merely makes it apparent that the real or objective situation is such that, in spite of the fact that there are no available parking spaces, the driver persists in trying to park his car. However, it is rather surprising that, in his analysis of the situation, Popper does not connect the irrationality of the driver to the *contradiction* between the information that the latter has at his disposal and the chosen course of action. For if the driver does not know or does not believe that the parking space where he is desperately trying to park his car is insufficient, then his behavior does *not* contradict the *RPs´* (Nadeau, 1993, pp. 461-62, emphasis added).

In an attempt to clarify this controversy Lagueux (2010, pp. 104f) notes that, the `flustered driver´ in Popper´s example, represents an atypical behaviour that cannot be excluded and that such behaviour can be said to be irrational. However, he recognizes that, `if we base our judgement *on the description alone*, we cannot be sure that the flustered driver´s action is really irrational´ (*op. cit*., p. 104). This is because we can *never* be certain that the agent sees P-S in a way that renders his behaviour appropriate. In other words, according to Lagueux (*op. cit*.) the same behaviour can be interpreted as either `rational´ or `irrational´ depending on the *observer´s viewpoint*. We thus believe Lagueux (*op. cit*.) ultimately admits that there is no objective or neutral way of deciding whether someone actually exhibits `means-rationality´ when there is no assurance that her beliefs are correct, i.e., when she does not exhibit `beliefs-rationality´. We conclude that, in the particular case of *RPs*, the dividing line between `rational´ and `irrational´ behaviour is unclear.

4.1.4. What version of *RP* is `preferable´ in the social sciences?

The theoretician´s methodological decision to ignore agents´ mistakes if *RPo* is adopted implies, for the reasons expounded above, assuming that agents´ (subjective) view of P-S *converges* over time to, or else, does not diverge in a significant way from the theoretician´s view of P-S. More specifically, we argued above that the adoption of *RPo* implicitly implies the *imposition* upon agents of the theoretician´s view of P-S. By contrast, and despite the above-mentioned tension between PTKL and *RPs*, we believe that *a more natural strategy in the theoretical social sciences is to adopt agents´ view of P-S*. There are at least three reasons for this. First, as Popper (1972, p. 179) admits, both the theoreticians´ and agents´ view of P-S are *conjectured*. Second, and more important, there is Hayek´s notion that, unlike the facts of the natural sciences — which are largely

independent of the theoretician´s viewpoint — the `facts´ of the social sciences are all *interpretations* (Hayek, 1943).[34] That is, according to Hayek, the concepts we use in the social sciences are not just abstractions like the ones used in the physical and natural sciences but they abstract from all the physical characteristics of the objects they refer to. Hayek (*op. cit*., p. 3) denotes the concepts we use in the theoretical social sciences as `teleological´ because, as he explains, such concepts can only be defined by postulating relations between three different terms: (i) a purpose, (ii) somebody who holds it, and (iii) an `object´ which the person in question thinks to be a suitable means to achieve that purpose. As he explains:

`We could say that all these objects are defined not in terms of their "real" properties but in terms of *opinions* people hold about them. In short, in the social sciences *the things are what people think they are*. Money is money, a word is a word, a cosmetic is a cosmetic, if and because somebody thinks they are. That this is not more obvious is due to the historical accident that in the world in which we live the knowledge of most people is approximately similar to our own… We are likely, for example, to think of the relationship between parent and child as an "objective" fact. But, when we use this concept in studying family life, what is relevant is not that *x* is the natural offspring of *y* but that either or both believe this to be the case´ (*op. cit*., emphasis added).

Third, Hayek (*op. cit*.) also argues that in the type of P-S analysed in the social sciences agents´ *interpretation* of P-S becomes an `integral´ part of the latter thereby affecting subsequent developments. In particular, to the extent that agents understand P-S via the internal models they create for that purpose, their understanding of the former will affect their decisions and, through this route, they *may affect P-S itself*. Let us use the example of `bank panics´ to illustrate this idea. The occurrence of a `bank panic´ in a private bank is not necessarily related to the *actual* liquidity position of the bank. Rather, the occurrence of a `bank panic´ is more likely to depend on its depositors´ view about the ability of the bank to cash their deposits on demand. If depositors have doubts about the ability of the bank to comply with its obligations when the former attempt to withdraw money from their accounts (and *regardless* of the `true´ liquidity position of

[34] This idea is discussed in detail also in Hayek (1942).

the bank), a `bank panic´ will likely ensue and the bank will *actually* become illiquid. This is not to deny that depositors commonly take into account the `objective´ indicators related to the liquidity of the bank when evaluating the likelihood of the private bank going illiquid. Rather, our argument is that *what really matters as far as depositors´ decisions are concerned is not the (objective) information provided by the liquidity indicators but agents´ (subjective) evaluation of them.* However, if the latter affects the former, then P-S is *not independent of depositors´ views* and, at least in this example, it is not sound to argue that there is an `objective´ P-S which is, in principle, knowable by the theoretician but *not* by the agents. This idea, we believe, is captured in the following comment by Hayek:

`Perhaps the relevant distinction comes out most clearly in the general and obvious statement that no *superior* knowledge the observer may possess about the object, but which is not possessed by the acting person, can help us in understanding the action in question´ (*op. cit*., emphasis added).

Unlike the presupposition by social scientists (and Popper) that the theoretician possesses a *wider* perspective of P-S than agents do, Hayek (*op. cit*.) suggests that, since P-S depends on agents´ interpretation no matter whether the latter is right or wrong, it follows that the theoretician does not stand in a privileged position to observe the `objective´ P-S. Thus, and for these reasons, we believe that the `natural´ strategy for the theoretician is to seek to capture P-S *as agents see it*. That is to say, if there is not an `objective´ P-S that is (fully) independent of agents´ views, then the theoretician has a better chance of understanding social phenomena if she adopts agents´ viewpoint. This is not to deny, however, that there may be some circumstances in which the theoretician may prefer instead, for methodological reasons, to adopt *RPo*. In particular, there may be circumstances where P-S may be sufficiently independent from agents´ beliefs as to make it convenient to adopt *RPo*. Be that as it may, the adoption of *RPo* will actually imply the *imposition* of the theoretician´s view of P-S upon agents or, as Schumpeter (1984) would put it, the model will capture the `rationality of the theoretician´ instead of the `rationality in the observed´.

Now, the former discussion suggests that *RPo* represents a *limit* or *extreme* case of *SA*. In particular, we believe *RPo represents a limit case of SA based on the implicit assumption that P-S is (fully) independent of agents´ beliefs* and that, consequently, the

theoretician can acquire `objective´ knowledge about P-S that is, somehow, *superior* to agents´. By contrast, Hayek´s ideas on the nature of the `facts´ of the social sciences imply that there are several elements of P-S such as the knowledge and information that agents possess and the social (and even physical) constraints their behaviour is subject to which depend on agents´ beliefs so the theoretician cannot claim to possess superior knowledge about them. In other words, the adoption of *RPo* could *a priori* be justified if the theoretician were able to acquire knowledge of P-S that is not available to agents but if, as Hayek (*op. cit*.) suggests, this is not the case, that is, if P-S consists, at least partly, of agents´ beliefs, if follows that *RPo* constitutes a *limit* case whose adoption implies *imposing* the theoreticians´ (allegedly superior) view of P-S on agents.

4.1.5. A reformulation of `situational analysis´

Popper apparently ignores both Hayek´s ideas about the peculiar `facts´ of the social sciences and his own ideas about indeterminism in the natural sciences in his discussion of *SA*. For instance, in his most detailed presentation of *SA* (Popper, 1994, p. 183, note 19), he argues that, if *RPo* is adopted, then the theorist reconstructs P-S *as it actually is* whereas, if *RPs* is adopted, she reconstructs P-S *as agents actually see it*. Yet, the way this is expressed by Popper is somewhat ambiguous since, in his attempt to clarify this issue, he seems to refer only to historical interpretation and, thus, it is unclear whether the distinction he draws between *RPo* and *RPs* also applies to the other social sciences. Hereafter, we assume that it does but we should like to make it clear that this is our interpretation. Specifically, he writes that:

> `It seems to me now that there are at least three senses of `rationality´ (and, accordingly, of the `rationality principle´), all objective, yet differing with regard to the objectivity of the situation in which the agent is acting: (1) *The situation as it actually was* — the objective situation which the historian tries to reconstruct. Part of this objective situation is (2) *The situation as the agent actually saw it*. But I suggest that there is a third sense intermediate between (1) and (2): (3) *The situation as the agent could* (within the objective situation) *have seen it*, and perhaps ought to have seen it´ (Popper, 1994, p. 183, footnote 19).

The previous quotation highlights that, when drawing a distinction between *RPo* and *RPs*, Popper assumes that the theoretician possesses knowledge that is superior to

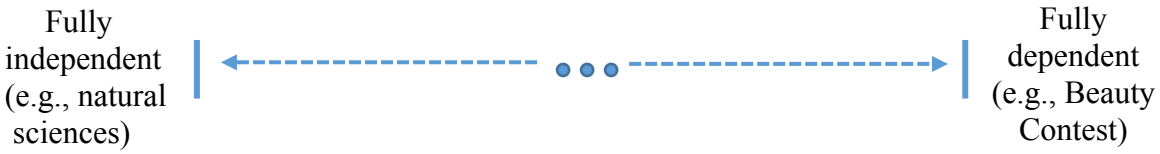
agents´. Otherwise, he could not have defined the third and intermediate sense of *RP* as one in which P-S is *`as the agent could* (within the objective situation) *have seen it*, and perhaps ought to have seen it´ (*op. cit.*). That is, the explicit reference by Popper to the P-S `as the agents *could* and perhaps *ought to* have seen it´ logically implies that he is assuming implicitly that there exists an `objective´ P-S and that the theoretician is in a *better* position than agents to observe it. Thus, we think Popper fails to take on board Hayek´s ideas about the peculiar `facts´ of the social sciences as well as his own views about indeterminism in the natural sciences. This takes us to our following claim. We believe that, arguably, the real difference between *RPo* and *RPs* is not that in the former the theorist reconstructs P-S *as it actually is* (even if understood as being conjectural) whereas in the latter she reconstructs it *as agents actually see it* (also understood in a conjectural way) but, instead, that in the former the theorist reconstructs P-S *as she sees it* whereas, in the latter, the theorist reconstructs it *as she believes agents see it*. This suggests that the difference between *RPo* and *RPs* in this reformulated framework is not the `objectivity´ of the approach — because both the theorist´s and agents´ view of P-S are subjective — but the *degree* in which the subjectivity of the theorist manifests itself; in the case of *RPo* the implied subjectivity of the theorist is of a `first degree´ because it is her (direct) view of P-S that is at stake whereas in the case of *RPs* the subjectivity of the theorist is of a `second degree´ because in that case it is her view about agents´ view of P-S.

Let us put it another way, if *RPo* is adopted the theorist reconstructs P-S *as she sees it* and, consequently, different theorists may reconstruct it in different ways (as it is commonly the case in the theoretical social sciences). Likewise, if *RPs* is adopted the theorist will reconstruct P-S *as she thinks agents see it*. Again, in this second scenario, different theorists may have very different views about *how* agents see P-S and, hence, may produce different theories about the same phenomenon. The reason is that, *if P-S is, at least partly, as agents see it, then it follows that Popper´s distinction between P-S `as it actually is´ and P-S `as agents see it´ makes little sense*. Indeed, we may add that in the extreme case in which P-S fully coincides with agents´ views (e.g., the Beauty Contest metaphor), *P-S is as agents see it* and, therefore, *RPo* and *RPs* would become equivalent if they were defined as Popper does.

Now, this suggests that, as we show in Chart 1 below, *SA* exhibits a spectrum of potential scenarios according to the *degree of independence* of P-S from agents´ beliefs. At one extreme of the spectrum there are those cases characterized by *full* coincidence

of P-S with agents´ beliefs or, as it were, by the absence of elements in P-S whose properties can be said to be (fully) independent from agents´ beliefs. Again, an example of this scenario is Keynes´ Beauty Contest metaphor in which there are no objective `facts´ the theoretician can observe because agents´ opinions about the (relative) beauty of the ladies portrayed in the photos are subjective. At the other extreme of the spectrum there is the typical scenario in the natural and physical sciences in which the `facts´ can be said to be *fully* independent from the observers´ viewpoint. Hayek´s (1943) suggests that in the social sciences there is no such scenario since P-S is *never* independent from agents´ beliefs and, hence, if we take Hayek´s ideas on board the former can be said to be a *limit* or extreme case. Between these two extreme cases there is a spectrum of potential P-S characterized by different positive degrees of dependence of P-S upon agents´ beliefs so that the lower the degree of dependence the closer the scenario will be to *RPo*. Finally, let us add that the case known as `self-fulfilling´ expectations (Merton, 1948) would correspond to the case in which agents´ beliefs, no matter whether they are right or wrong, bring about a change in P-S so that the latter eventually *converges* to the former. In our framework, this could only occur if P-S were at one end of the above-mentioned spectrum; the one characterised by coincidence of P-S with agents´ beliefs.

Chart 1. Spectrum of scenarios according to the degree of `independence´ of P-S from agents´ beliefs



Let us finish off this section by adding that the former discussion highlights that, if *RPo* is adopted, the null hypothesis in an empirical test is not that the theorist´s view of the typical P-S is correct, as it is usually believed, but rather that *agents´ view of the typical P-S coincides, on average, with the theorist´s*. That is to say, to the extent that agents´ view of P-S is an integral part of the latter, only if there is a large coincidence (on average) between the agents´ and the theorist´ view of P-S will the hypothesis have a chance of withstanding the *onus of proof* when subject to an empirical test. In other words, an empirical test is not, if *RPo* is adopted, a contrast between the theorist´s view of P-S and the `objective´ facts, as is the case in the natural sciences. Rather, as we have

argued above, the `facts´ of the social sciences are largely *interpretations* and, hence, rejection of the null hypothesis in this case does not imply that the theorist´s view of P-S does not capture the `objective´ facts appropriately since there are no `objective´ facts in the social sciences but, rather, that the theorist´s view of P-S does not coincide, on average, with agents´ view of P-S.

4.2. PTKL *versus SA*: the `rationality of the theoretician´

The last issue we should like to address is the compatibility or otherwise of the notion of rationality that stems from PTKL and *SA* as viewed from the perspective of the theoretician. If the discussion in the previous section has focused on the rationality of the agents who are the object of modelling by the theoretician, this section focuses on the `rationality of the theoretician´ or, as Schumpeter (1984) denotes it, the `rationality of the observer´. We believe that the apparently irreconcilable notions of rationality that stem from PTKL and *SA* when looked at from the standpoint of the agents *consist, when approached from the standpoint of the theoretician, of the application of PTKL to two different problems*. First, we share Lagueux´s argument that `true´ rationality in Popper actually consists of the corrigibility of one´s beliefs (Lagueux, 2006, p. 202).[35] Second, and according to Popper, science represents a particular version of PTKL characterized by the application of the `critical method´:

`The difference between the amoeba and Einstein is that, although both make use of the method of trial and error or elimination, the amoeba dislikes erring while Einstein is intrigued by it: he consciously searches for his errors in the hope of learning by their discovery and elimination. The method of science is the critical method´ (Popper, 1972, p. 70).

Now, if science is characterized as the application of the `critical method´ to the object of knowledge, *we may rationalize SA as the specific application of the ´critical method´ to the social sciences*. That is, the theoretician of the social sciences proposes a reconstruction of the P-S in which actors find themselves by formulating a conjectural

[35] For instance, in *The Open Society and its Enemies*, Popper ([1943*a*] 1966, p. 97) states that `the method of applying a situational logic to the social sciences is not based on any psychological assumption concerning the rationality (or otherwise) of "human nature"´. In this respect, Oakley (1999, p. 35) notes that Popper stresses in his Harvard lecture (Popper, 1985, 1994) that, in the context of *SA*, `rationality had no ontological significance and was not intended as a theory of human action´.

situational model where they make decisions on the basis of systematic deliberation. As such, the situational model proposed is only a conjecture according to which the former constitutes an `adequate´ oversimplification of the relevant P-S or:

`By conjectural analysis I mean a certain kind of tentative or conjectural explanation of some human action which appeals to the situation in which the agent finds himself… Admittedly, no creative action can ever be fully explained. Nevertheless, we can try, conjecturally, to give an idealized reconstruction of the problem situation in which the agent found himself, and to that extent make the action "understandable" (or "rationally understandable")…´ (Popper, 1972, p. 179).

The situational model constructed will then be subject to close scrutiny by the scientific community which will tell us the extent to which the different elements of the hypothetical P-S need to be modified. To the extent that the elements of the hypothetical P-S can be properly identified (i.e., it should consist of a set of observable elements such as physical and social constraints and assumptions about agents´ knowledge and information) both the situational model and the results of the empirical tests can be, in principle, openly criticised by the scientific community. That is, the significance of the situational model is that it is the `object´ against which criticism may be directed given the prior methodological decision to immunize *RP* from potential refutation. In turn, the role of *RP* (regardless of the specific version adopted) is to facilitate the implementation of the `critical method´ by helping scientists identify the `logic of the situation´ captured in the situational model. In this respect, it has been argued elsewhere that, in the context of Popper´s distinction among three ontological domains, the claim to *objectivity* in World 3 — which consists of knowledge or thought in an objective sense such as problems, theories, and arguments — stems from the notion that knowledge in World 3 `resides in *recorded* form outside the mind of any agent, even its originator, and is, in principle, accessible by any other agent in that recorded form´ (Oakley, 2002, p. 464; also Popper, 1972, pp. 108-9). As Sassower (2006, p. 104) has observed, Popper `saw rationality as the way to intersubjectivity, because it is too much to expect objectivity´. In the absence of an `animating principle´ such as *RP*, it would be very hard to logically connect the elements of the situational model in an understandable way.

What is crucial, however, is that the situational model will undergo successive changes and refinements (that will let it acquire greater accuracy in the explanation of

social phenomena) in the aftermath of `rational´ criticism by the scientific community. Viewed from this standpoint, it does not make any difference whether the theoretician adopts *RPo* or *RPs* as long as the situational model is subject to close scrutiny by other members of the scientific community. Then, as the situational model undergoes further refinements or modifications new predictions and explanations may eventually emerge. It is clear then that, as long as social scientists modify their models in order to make them capable of providing increasingly accurate explanations of social phenomena, *their behaviour can be characterized as being `corrigible´ and, hence, as sticking to the type of rationality we have associated with PTKL*. We may thus conclude that, although there is some tension between PTKL and *SA* when their relation is approached from the standpoint of `the rationality of agents´ such tension does not arise when such relation is approached instead from the standpoint of the `rationality of the theoretician´.

5. Summary and conclusions

A number of commentators have noted that there are two different approaches to rationality in Popper´s philosophy: the approach stemming from his evolutionary theory of knowledge and learning (PTKL) and the approach embodied in so-called `Situational Analysis´ (*SA*) and associated to his famous `Rationality Principle´ (*RP*). According to the former, we `learn´ by subjecting our hypotheses to trial and discarding those ones which turn out to be wrong. In addition, all knowledge is conjectural and fallible. In this setting, science is the `highest´ form of knowledge acquisition and is characterized by the subjecting of scientific theories to the most severe forms of criticism by members of the scientific community including, of course, empirical testing. The notion of `critical rationalism´ was coined to capture Popper´s thesis that the way to maximize the rate of expansion of knowledge is to subject theories to an optimal amount of criticism. The notion of human rationality that stems from PTKL consists of the *corrigibility* of our (wrong) beliefs. In other words, PTKL implies that being `rational´ consists of revising our beliefs when they turn out to be wrong. Therefore, individuals who fail to do so are `irrational´. *SA* represents Popper´s methodological proposal for the social sciences. In turn, *RP* is a methodological principle according to which *agents always act in a way that is adequate or appropriate to their problem-situation* (P-S). It follows that, in the context of *SA*, `rational´ behaviour consists of acting in a way that is appropriate to the `logic of the situation´ whereas `irrational´ behaviour will consist of doing otherwise. However, Popper and several of his commentators have made an important distinction

between the `objectivist´ and the `subjectivist´ version of *SA*. According to them, in the former the theoretician seeks to reconstruct P-S `as it actually is´ whereas in the latter she reconstructs it `as it is seen by agents´. In this respect, we argued that the former is based on the (implicit) assumption that there is a systematic *discrepancy* between the theoretician´s and agents´ view of P-S and that, under the null hypothesis, this implies that agents´ view of P-S is assumed to be, at least partially, wrong. We also proposed a rationalization of the `objectivist´ version of *SA* according to which the latter is based on a methodological decision to assume that the mistakes made by agents when making decisions are *less* interesting for the purpose of understanding agents´ behaviour and, especially, for the generation of predictions than the (modelling) mistakes made by the theoretician.

The purpose of this essay was to study the compatibility of these two apparently irreconcilable approaches. We have made five different claims. Our first claim was that there is a certain tension between PTKL and *SA* when their relation is analysed from the standpoint of the `rationality of the agents´ whose behaviour is captured in the model albeit the tension disappears when the relation is analysed from the standpoint of the `rationality of the theoretician´. Our second claim was that the nature of the tension between PTKL and *SA* depends on whether the theoretician adopts the `objectivist´ or the `subjectivist´ version of *SA*. In particular, we argued that the tension between PTKL and the `subjectivist´ *SA* stems from the fact that, in the latter, it is implicitly assumed that agents´ view of P-S is partially wrong which implies, in turn, that agents do not `learn´ from their mistakes as PTKL posits, i.e., that they exhibit a tendency to repeat the same mistakes so that the latter become predictable. We argued that this feature of the `subjectivist´ version of *SA* creates a tension with PTKL when the main purpose of the theoretician is to generate predictions but does not generate any tension with PTKL when the purpose is to perform historical interpretation. This raises the difficult issue of the legitimacy of adopting the `subjectivist´ *SA* when the main purpose of the modelling exercise is to generate predictions the latter being understood as the derivation of the logical consequences of theory. It was not the purpose of this study to settle this issue and, hence, issuing a verdict on it will require further work. By contrast, we argued that the tension between PTKL and the `objectivist´ *SA* stems from the fact: (i) that if agents behave according to PTKL it is not necessarily the case that their decisions will be adequate or appropriate to the `logic of the situation´ insofar as the former *only* implies that agents tend to eliminate their mistakes and, hence, we have that in the wake of

changes in the surrounding environment agents´ decisions may not be adequate to the `logic of the (new) situation´, and (ii) that the adoption of the `objectivist´ *SA* implies *de facto* the *imposition* of the theoretician´s view of P-S upon agents´. Yet, we argued that it is unlikely that, if agents behave according to PTKL by subjecting their conjectures to trial and eliminating the ones that turn out to be wrong, their view of P-S will eventually converge to the theoretician´s. Our third claim built on the ideas of Hayek (1943) about the nature of the `facts´ of the social sciences and was that, in the way it is presented by Popper and some of his commentators, the `objectivist´ *SA* represents a *limit* or extreme case based on the presupposition that P-S is (fully) *independent* of agents´ beliefs. Our fourth claim was closely related to the previous one and consisted of the idea that, if Hayek´s ideas on the nature of the `facts´ of the social sciences are duly taken on board, it follows that the natural strategy for social scientists is to seek to reconstruct P-S *as agents´ see it* rather than to reconstruct it *as the scientist sees it*. Our fifth and last claim was that, unlike what Popper and his commentators suggest, the difference between the `objectivist´ and the `subjectivist´ version of *SA* is not that in the former the theoretician reconstructs P-S *as it actually is* whereas in the latter she does it as *agents see it* but, rather, that in the former she reconstructs P-S *as she sees it* herself whereas in the latter she does it *as she believes that agents actually see it*.

References

Alchian, A. A. (1950) Uncertainty, Evolution, and Economic Theory, *Journal of Political Economy*, 58(3), pp. 211-221.

Arrow, K. J. (1986) Rationality of self and others in an economic system, *Journal of Business*, 59(4.2): S385-S399.

Arrow, K. J. (2004) Is bounded rationality unboundedly rational? Some ruminations, in Augier, M. & March, J. G. (eds.) *Models of a Man: Essays in Memory of Herbert A. Simon*, Cambridge, MA: MIT Press, pp. 47–55.

Bartley, W. W. (1982) The Philosophy of Karl Popper. Part III. Rationality, Criticism, and Logic, *Philosophia* 11(1–2), pp. 121–221.

Becker, G. (1962) Irrational behavior and economic theory, *Journal of Political Economy*, 70(1), pp. 1-13.

Becker, G. S. (1976) *The economic approach to human behaviour,* Chicago (IL): University of Chicago Press.

Beinhocker, E. D. (2013) Reflexivity, complexity, and the nature of social science. *Journal of Economic Methodology*, 20(4), pp. 330-342.

Bentham ([1789] 1907) *An Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press

Bicchieri, C. (1992) Two Kinds of Rationality, in de Marchi, N. (ed.) *Post-Popperian Methodology of Economics: Recovering Practice*, pp. 155-188, Boston: Kluwer Academic Publishers.

Blaug, M. (1975) Kuhn versus Lakatos, or Paradigms versus Research Programmes in the History of Economics, *History of Political Economy*, 7, pp. 399–419.

Blaug, M. (1985) Comment on D. Hands, 'Karl Popper and economic methodology: a new look', *Economics and Philosophy*, 1(2), pp. 286–8.

Boland, L. A. (2003*a*) Dealing with Popper in Economic Methodology, *Philosophy of the Social Sciences,* 33(4), pp. 479-98.

Boland, L. A. (2003*b*) *The Foundations of Economic Method: A Popperian Perspective*, 2nd edition, London: Routledge.

Bunge, M. (1996) The Seven Pillars of Popper´s Social Philosophy, *Philosophy of the Social Sciences*, 26(4), pp. 528-556.

Caldwell, B. J. (1985) Towards a broader conception of criticism, *History of Political Economy*, 18(4), pp. 675-81.

Caldwell, B. J. (1991) Clarifying Popper, *Journal of Economic Literature*, 29(1), pp. 1-33.

Caldwell, B. J. (2003) Popper and Hayek: Who Influenced Whom? *Cashiers D´Épistémologie*, 292, 2003-01.

Campbell, D. T. (1974) Evolutionary Epistemology, in Schilpp, P. A. (ed.) *The Philosophy of Karl Popper*, pp. 413-463, Vol. XIV, Book I, LaSalle (IL): Open Court.

Cosmides, L. & Tooby, J. (1994) Better than Rational: Evolutionary Psychology and the Invisible Hand, *American Economic Review*, 84(2), pp. 327-32.

De Bruin, B. (2006) Popper´s Conception of the Rationality Principle in the Social Sciences, in Jarvie, I., Minford, K. & Miller, D. (eds.) *Karl Popper: A Centenary Assessment*, Vol. III (Science), pp. 209-217, Ashgate.

Dow, S. C. (1995) The appeal of neoclassical economics: some insights from Keynes's epistemology, *Cambridge Journal of Economics*, 19(6), pp. 715-733.

Edgeworth (1881) *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. Augustus M. Kelley: New York.

Farmer, M. K. (1982) Rational Action in Economic and Social Theory: Some Misunderstandings, *Archives Europeennes de Sociologie*, 23(1), pp. 179-197.

Farmer, J. D. (1998) *Market force, ecology and evolution*. Santa Fe Institute Working Paper.

Friedman, M. (1953) The methodology of positive economics, in *Essays in Positive Economics*, ch. 1, pp. 3-43, Chicago: Chicago University Press.

Gattei (2009) *Karl Popper's Philosophy of Science,* Routledge: New York.

Gibbard, A. & Varian, H. R. (1978) Economic Models, *Journal of Philosophy*, 75(11), pp. 664-677.

Goldthorpe, J. H. (1998) Rational Action for Sociology, *British Journal of Sociology*, 49(2), pp. 167-192.

Hamlin, A. P (1986) *Ethics, Economics and the State*, New York: St. Martin´s Press. Reprinted in Caldwell, B. (ed.), 1993, *The Philosophy and Methodology of Economics I*, pp. 304-362, Aldershot: Edward Elgar.

Hands, D. W. (1985) Karl Popper and Economic Methodology: A New Look, *Economics & Philosophy*, 1(1), pp. 83-99.

Hands, D. W. (1991) Popper, the rationality principle and economic explanation, in G. K. Shaw (ed.) *Economics, Culture and Education: Essays in Honour of Mark Blaug*, Cheltenham: Edward Elgar, pp. 108-119.

Hands, D. W. (1992) Falsification, Situational Analysis and Scientific Research Programs: The Popperian Tradition in Economic Methodology, in de Marchi, N. (ed.) *Post-Popperian Methodology of Economics: Recovering Practice*, pp. 19-54, Boston: Kluwer Academic Publishers.

Hayek, F. A. (1937) Economics and Knowledge, *Economica*, 4(13), pp. 33-54.

Hayek, F. A. (1942) Scientism and the Study of Society. Part I, *Economica*, 9(35), pp. 267-291.

Hayek, F. A. (1943) The Facts of the Social Sciences, *Ethics*, 54(1), pp. 1-13.

Hayek, F. A. (1967) *Studies in Philosophy, Politics, and Economics,* Chicago: University of Chicago Press.

Hedström, P., Swedberg, R. & Udéhn, L. (1998) Popper´s Situational Analysis and Contemporary Sociology, *Philosophy of the Social Sciences*, 28(3), pp. 339-364.

Hicks, J. R. (1956) *A Revision of Demand Theory*, Oxford: Clarendon Press.

Hobbes, T. ([1651] 1985) *Leviathan*, London: Penguin Books.

Hume, D. ([1748] 2006) An enquiry concerning human understanding: a critical edition, Oxford, Clarendon Press; New York : Oxford University Press.

Jacobs, S. (1990) Popper, Weber and the Rationalist Approach to Social Explanation, *British Journal of Sociology*, 41(4), pp. 559-570.

Kerstenetzky, C. L. (2009) Plural situational logic: the rationa(lisabi)lity principle, *Cambrigde Journal of Economics*, 33(2), pp. 193-209.

Keynes, J. M. (1936) *The General Theory of Employment, Interest and Money.* Cambridge: Macmillan.

Kirzner, I. (1976) *The Economic Point of View*, Kansas City: Sheed & Ward.

Klappholz, K. & Agassi, J. (1959) Methodological Prescriptions in Economics, *Economica*, 26(101), February, pp. 60-74.

Knight, F. H. ([1921] 1971) *Risk, Uncertainty, and Profit*, Chicago: The University of Chicago Press.

Koertge, N. (1975) Popper's Metaphysical Research Program for the Human Sciences, *Inquiry*, 18(4), pp. 437-62.

Koertge, N. (1979) The Methodological Status of Popper´s Rationality Principle, *Theory and Decision*, 10(1/4), January, pp. 83-95.

Lagueux, M. (1993) Popper and the Rationality Principle, *Philosophy of the Social Sciences*, 23(4), pp. 468-480.

Lagueux, M. (2006) Popper and the Rationality Principle, in Jarvie, I., Minford, K. & Miller, D. (eds.) *Karl Popper: A Centenary Assessment*, Vol. III (Science), pp. 197-217, Ashgate.

Lagueux, M. (2010) *Rationality and Explanation in Economics*, Routledge.

Langlois, R. N. (1995) Rationality, institutions, and explanation, in R. N. Langlois (ed.) *Economics as a Process*, pp. 225-255, Cambridge: Cambridge University Press.

Langlois, R. N. (1998) Rule-following, Expertise, and Rationality: A New Behavioral Economics? in Dennis, K. (ed.) *Rationality in Economics: Alternative Perspectives*, pp. 55-78, Boston: Kluwer Academic.

Latsis, S. J. (1972) Situational Determinism in Economics, *British Journal for the Philosophy of Sciences*, 23(3), pp. 207-245.

Latsis, S. J. (1983) The Role and Status of the Rationality Principle in the Social Sciences, in Cohen, R. S. & Wartoksky, M. W. (eds.) *Epistemology, Methodology, and the Social Sciences*, pp. 123-152, London: Reidel Publishing Company.

Loasby, B. J. (1976) *Choice, complexity and Ignorance: An Enquiry into Economic Theory and the Practice of Decision-Making*, Cambridge: Cambridge University Press.

Loasby, B. J. (1999) *Knowledge, Institutions and Evolution in Economics*, London: Routledge.

Lucas, R. E. (1986) Adaptive behaviour and economic theory, in R. M. Hogarth & M. W. Reder (eds.) *Rational Choice: The Contrast between Economics and Psychology*, pp. 217-242, Chicago and London: University of Chicago Press.

Matzner, E. & Jarvie, I. C. (1998) Introduction to the Special Issues on Situational Analysis, *Philosophy of the Social Sciences*, 28(3), pp. 333-338.

Mayr, E. (1988) *Toward a New Philosophy of Biology – Observations of an Evolutionist.* Cambridge, Mass., and London: Harvard University Press.

Menger, C. ([1871] 1950) *Principles of Economics,* translated and edited by J. Dingwall & Bert H. Hoslitz. Glencoe, IL: Free Press.

Merton, R. K. (1948) The self-fulfilling prophecy, *Antioch Review,* 8(2), pp. 193-210.

Mill, J. S. ([1836] 1967) On the Definition of Political Economy and the Method of Investigation Proper to It, in *Collected Works of John Stuart Mill*, vol. 4, pp. 309-39, Toronto: University of Toronto Press.

Minsky, H. P. (1975) *John Maynard Keynes*. New York: Columbia University Press.

Muth, J. F. (1961) Rational Expectations and the Theory of Price Movements, *Econometrica*, 29(3), pp. 315-335.

Nadeau, R. (1993) Confuting Popper on the Rationality Principle, *Philosophy of the Social Sciences*, 23(4), pp. 446-467.

Notturno, M. A. (1998) Truth, Rationality, and the Situation, *Philosophy of the Social Sciences*, 28(3), pp. 400-421.

Oakley, A. (1999) Economics and the origin of Popper´s situational analysis, *History of Economics Review*, 30, pp. 25-40.

Oakley, A. (2002) Popper´s Ontology of Situated Human Action, *Philosophy of the Social Sciences*, 32(4), pp. 455-486.

Pareto, V. (1917) *Traitè de sociologie generale*, Payot: Lausanne and Paris.

Parsons, T. (1937) *The Structure of Social Action*, McGraw Hill: New York.

Popper, K. R. ([1943*a*] 1966) *The Open Society and its Enemies*, Vol. II, London: Poutledge & Kegan.

Popper, K. R. ([1943*b*] 1966) *The Open Society and its Enemies*, Vol. I, London: Poutledge & Kegan.

Popper, K. R. (1944-45) The Poverty of Historicism I-III, *Economica*, 11, pp. 86-103, 119-137, 12: pp. 69-89.

Popper, K. R. (1963) *Conjectures and Refutations: The Growth of Scientific Knowledge*, London: Routledge and Kegan Paul.

Popper, K. R. (1966) *Of Clouds and Clocks: An Approach to the Problem of Rationality of and the Freedom of Man*, St. Louis: Washington University Press.

Popper, K. R. (1967) La rationalité et le statut du principe de rationalité, in E. M. Classen (ed.), *Les Fondements Philosophiques des Systèmes Economiques*, pp. 142-50, Paris: Payot.

Popper, K. R. (1972) *Objective Knowledge: An Evolutionary Approach*, Oxford: Clarendon Press.

Popper, K. R. (1976*a*) *Unended Quest*, LaSalle, IL.: Open Court.

Popper, K. R. (1976*b*) The Logic of the Social Sciences, in T. W. Adorno et al., *The Positivist Dispute in German Sociology*, pp. 87-104, London: Heinemann.

Popper, K. R. (1985) The Rationality Principle, in D. Miller (ed.) *Popper Selections*, pp. 357-365, Princeton: Princeton University Press.

Popper, K. R. (1990) *A World of Propensities*, Bristol (UK): Thoemmes.

Popper, K. R. (1994) *The Myth of the Framework: In defence of science and rationality*, London: Routledge.

Sassower, R. (2006) *Popper´s Legacy: Rethinking Politics, Economics and Science*, Stocksfield (UK): Acumen.

Savage, L. (1954) *The Foundations of Statistics,* New York: Wiley.

Schumpeter, J. A. (1984) The Meaning of Rationality in the Social Sciences, *Journal of Institutional and Theoretical Economics*, 140(4), pp. 577-593.

Schutz, A. (1972) *The Phenomenology of the Social World*, London: Heinnemann.

Simon, H. A. (1976) From Substantive to Procedural Rationality, in Spiro J. Latsis (ed.) *Method and Appraisal in Economics*, pp. 129-148, Cambrigde: Cambridge University Press.

Twomey, P. (1998) Reviving Veblenian Psychology, *Cambridge Journal of Economics,* 22(4), pp. 433-48.

Tversky, A. & Kahneman, D. (1986) Rational choices and the Framing of Decisions, in R. M. Hogarth & M. W. Reder (eds.) *Rational Choice: The Contrast between Economics and Psychology*, pp. 67-94, Chicago: University of Chicago Press.

Vanberg, V. J. (2002) Rational Choice vs. Program-based Behaviour: Alternative Theoretical Approaches and their Relevance for the Study of Institutions, *Rationality and Society*, 14(7), pp. 7-54.

Volz, K. H. & Gigerenzer. G. (2012) Cognitive processes in decisions under risk are not the same as in decisions under uncertainty, *Frontiers in Neuroscience*, 6(1), pp. 1-6.

von Mises, L. (1944) The Treatment of "Irrationality" in the Social Sciences, *Philosophy and Phenomenology Research*, 4(4), pp. 527-546.

von Neumann, J. & Morgenstern, O. (1947) *Theory of Games and Economic Behaviour*, Princeton (NJ): Princeton University Press.

Vromen, J. J. (1995) *Economic Evolution: An Enquiry into the Foundations of New Institutional Economics*, London: Routledge.

Weber, M. (1904-5) *The Protestant Ethic and the Spirit of Capitalism*, London: Allen & Unwin.

Weber, M. (1949) *The Methodology of the Social Sciences*, New York: Free Press.

Wible, J. R. (1984/85) An epistemic critique of rational expectations and the neoclassical macroeconomics research program, *Journal of Post-Keynesian Economics*, 7(2), pp. 269-81.

Zouboulakis, M. S. (2014) *The Varieties of Economic Rationality: From Adam Smith to contemporary behavioural and evolutionary economics*, Routledge.