The logic of scientific debate: Epistemological quality control practices and Bayesian inference – a neoPopperian perspective

John R. Skoyles CPNSS
j.skoyles@lse.ac.uk
www.human-existence.com

Science is about evaluation, persuasion and logic. In scientific debate, scientists collectively evaluate theories by persuading each other in regard to epistemological qualities such as deduction and fact. There is, however, a flaw intrinsic to evaluation-by-persuasion: an individual can attempt and even succeed in persuading others by asserting that their reasoning is logical when it is not. This is a problem since, from an epistemological perspective, it is not always transparent nor obvious when a persuasive assertion is actually deductively warranted. Empirical research upon reasoning, indeed, supports the notion that assertions are often strongly persuasive for reasons other than their logic.

The unreliability of the link between persuasion and logic raises an important methodological issue: how do scientists debate in a manner such that claimed but false "logical" arguments are ignored, and only warranted arguments get to determine theory preference? This need for soundness in debate is a particularly important epistemological concern in cases where the deductive qualities of persuasive argument are not overt, and so cannot be directly checked --such as when they are founded upon Bayesianism probabilistic coherence.

The argument presented here is that scientists make the qualities of probabilistic and nonprobabilistic inference sound (and so warranted) through how they organize their debate. Scientists, I argue, abide by "epistemological quality control practices" that limit the persuasive power of unsound arguments upon theory evaluation. Examples of such debate quality control practices are publicness, clarity, openness to criticism, and the collective promotion of attempts at theory conjecture and refutation.

Methodologically, these quality control practices are extralogical since they do not directly provide scientific inferences with additional logical warrantedness. They function instead in science to generate an epistemological evaluative environment in which persuasiveness is due, and only due, to logic (i.e. sound). Their methodological role is therefore to make what is warranted *de papyri* (in our principles of rationality – epistemologically competence) also what is persuasive in evaluation and debate *de cognitio* (in our exercise of such principles – epistemologically performance).

Several limitations exist upon the soundness of Bayesian inferential coherence – surreptitious revision, logical omniscience, uncertain evidence, old evidence, and new hypotheses. Bayesianism, as a result, can only exist if it is pursued in a debate that is regulated by quality control practices over its inferences (for instance, practices that ensure there is autonomy of inference, diligence of inference, probity of evidence, auditability of inference and assiduousness of conjecture). In this context, one can reinterpret Popper's concern with criticism, openness, refutation and conjecture, as deriving not (as he thought) directly from the needs of logic, but, indirectly, from the need of scientists to create epistemological soundness.

Contents

1 Introduction

1.1 Scientific debate and scientific inference

Philosophy of science has a number of established and recognized concerns such as examining the basis of rational theory preference and the nature of explanation. Here this paper introduces another topic for investigation: the epistemological importance of the institutions, practices and ethos that regulate the logical soundness of the debate in which scientists evaluate theories.

Intuitively, persuasion amongst scientists should be sound – that is done – and only done – in regard to the warrantedness provided by logical deduction combined with empirical fact.

But things are not so epistemologically simple. For example, (i) our facts are often contaminated by artefacts, inadequate controls, and misunderstood statistics; and (ii) our cognitive attempts at deduction are often compromised by prior convictions, confirmation biases, bounded rationality, self-interest, deference, and anxieties about how the outcome of a theory evaluation might determine personal career advancement.

As a result, empirical and deductive claims that are treated and considered in scientific evaluation to be valid, might often seem to present these qualities to scientists when, in fact, they lack empirical or deductive warrantedness.

Scientific evaluation, however, if it is to be epistemologically rational, must be exclusively based on logically warranted assertions.

This raises the question of how scientists can engage in epistemologically sound debate when making preferences between theories. Can institutional and other practices be added to scientific debate that act to prevent it from being corrupted by the presence of assertions that persuade in regard to false logicalness or warrantedness? This piece explores this methodological problem, and argues for the proposal that scientists have developed alongside experiments and theoretical analysis, methodological procedures that function to prevent unsound theory evaluation. These procedures will be called here "epistemological quality control practices". These practices are embedded in the institutions, intellectual discourse and ethos of science.

1.2 De papyri and de cognitio logic levels

One approach to articulate the above problem as a methodological issue is to distinguish between the existence of two epistemological levels (or realms) in which the logical evaluation of theories in scientific debate can be said to occur: *de papyri* (on paper or in theory), and *de cognitio* (through cognition, mental process, or mind). While this distinction is rarely made, there is much utility for those examining science and its reasoning to use it to formulate the nature of the process and the logical foundations of scientific debate (see also fig. 1 and accompanying legend for these and related concepts).

The realm of *de papyri* logic directly concerns deductions and reasons: its epistemological subject matter is that of the processes of the abstract and unembodied rationality explored by philosophers and logicians – the qualities of coherence and logic of propositions that exist only in theory and on paper. The realm of *de cognitio* logic concerns, instead, with what epistemologically occurs at the level of human minds as they think as deducers and reasoners of such propositions. It concern is with propositions as they are epistemologically embodied in the actual attempts of scientists to be rational in their cognitions, statements and assertions. The critical point here is that entities at the level of *de papyri* logic never themselves argue, believe, persuade, debate nor evaluate. These activities are restricted to the people that seek and claim to use them to be rational. The principles of logic, however, are just that "principles". As such they never epistemologically do anything by themselves. It is only at the level of *de cognitio* logic that arguments gain the epistemological quality of being claimed to be valid and warranted, and so effect the evaluation of scientific theories.

Another way of making this epistemological distinction is to note its parallel with the Chomskyan language competence/ performance difference. *De papyri* logic in the context of inferences about external realities is concerned with epistemological competences, while *de cognitio* reasoning about empirical data is concerned with epistemological performance. Epistemological competence provides a rational account of what can be inferred about reality that is abstract from any actual individual or community engaging in such reasoning. Epistemological performance concerns what can be rationally done by individuals or communities of individuals that seek to do so such reasoning.

The epistemological divide noted here between the realms of *de cognitio* and *de papyri* logic, and that between epistemological competence and epistemological performance, raises the methodological question as to how they could be linked. How can science done by humans at a *de cognitio* level make itself rational at a *de papyri* one? How can the epistemological performance of scientific reasoning link to the epistemological competence of making valid truth evaluating inferences?

Specifically, how is it possible for a community of reasoners, "scientists", when faced with an argument that claims logical warrantedness, to epistemologically know that that its asserted argument possesses (in spite of this not being necessary), the underlying logical inference and deductive coherence warrantedness quality it claims. This requirement – that claimed inferences must be real and not "logical" only in appearances, I shall call the *actuality requirement*.

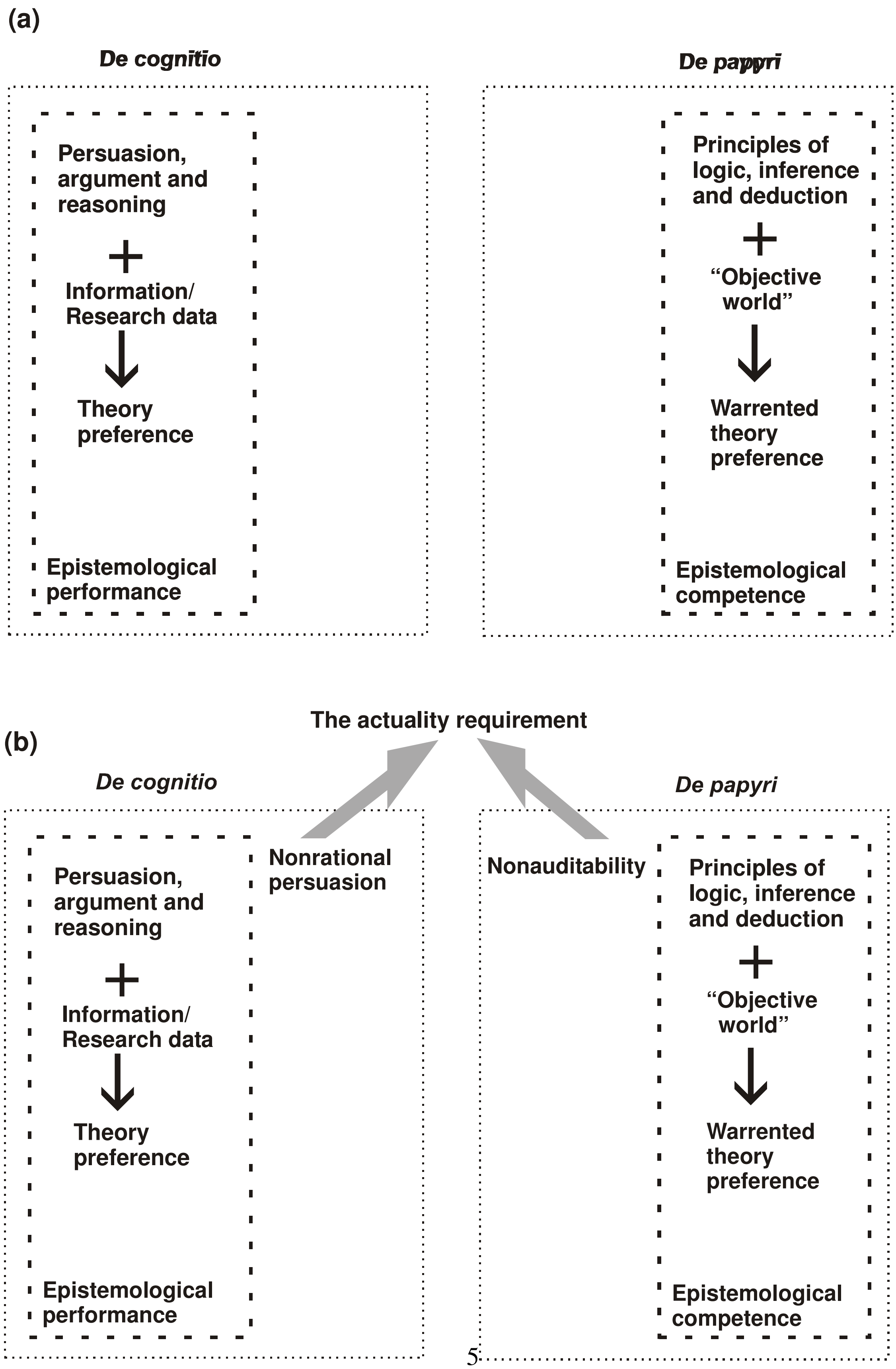
One solution – and it is one that this paper proposes that scientists use – is to regulate debate with practices that ensure that the actuality requirement is satisfied. Such practices allow scientists to minimize the presence amongst themselves of reasoning that is epistemologically unsound – that is reasoning that while it might be valid is unwarranted (it might be logically valid or then again it might not).

This raises the methodological question as to the nature of the practices that scientists might adopt as a community to ensure that their debate and reasoning from an methodological viewpoint is epistemologically sound and fulfils the actuality requirement.

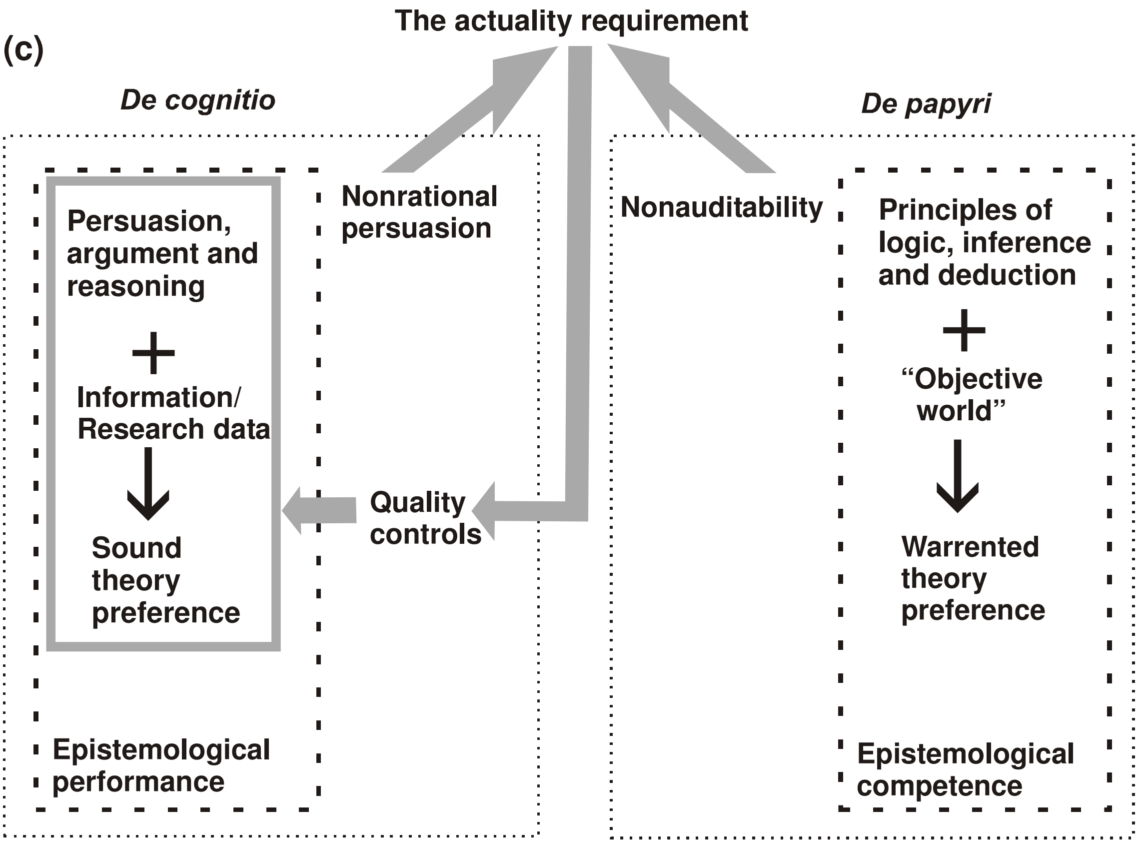
1.3 The actuality requirement and epistemological quality control practices

The argument of this paper is first, that for the scientific community, solving the problem of the actuality requirement and so the soundness of evaluative debate is a nontrivial problem. Second, as a result, a core part of science is the existence of a "reasoning craft" of traditions and standards used by scientists that minimize, if not eliminate over the long-term, the presence of false persuasiveness when they debate theories. Indeed, if we look at scientists when they reason, they can be seen to use methodological practices that have little direct connection with logic but which seem only to exist to ensure that their reasoning is persuasive, and only persuasive, due to logic.

Figure 1


(a)
De cognitio
Persuasion, argument and reasoning
+
Information/ Research data
Theory preference
Epistemological performance
De papyri
Principles of logic, inference and deduction
+
“Objective world”
Warrented theory preference
Epistemological competence
(b)
The actuality requirement
De cognitio
Persuasion, argument and reasoning
+
Information/ Research data
Theory preference
Epistemological performance
Nonrational persuasion
De papyri
Nonauditability
Principles of logic, inference and deduction
+
“Objective world”
Warrented theory preference
Epistemological competence

(c)



In figure 1 (a), the situations of *de cognitio* and *de papyri* realms is illustrated. Within the *de cognitio* realm exists the persuasions, arguments and reasoning that people make with the information and research data used to prefer theories. This makes up epistemological performance. Within the *de papyri* realm exists the principles of logic, inference and deduction that can be used to make warranted theories about the objective world. This makes up epistemological competence.

In figure 1 (b), the situation is shown to be more complex. In the *de cognitio* realm exists nonrational persuasion, and in the *de papyri*, nonauditability which creates the need for the actuality requirement that links these two realms of reasoning.

In figure 1 (c), the solution offered in this paper is shown. The actuality requirement creates the need for quality controls upon epistemological performance. In doing this it turns its theory preferences in the *de cognitio* realm into epistemologically sound theory preferences.

This paper is innovative in that those presently studying the methodology of scientific reasoning rarely discuss the distinction between *de papyri* and *de cognitio* epistemological reasoning, epistemological competence and epistemological performance, the existence of the actuality requirement, evaluative soundness, and so the critical methodological role and importance of epistemological quality control practices in scientific reasoning and debate.

As a result, these philosophical issues have failed to be given their due importance in the study of the foundations of science as a collective rational activity. The methodology of science as a result has focused entirely upon how scientific reasoning reflects logical warrantedness without covering how it also enables scientific theory preferences to be logically sound.

2 De papyri and de cognitio levels and bayesianism

Bayesian probability theory provides the dominant model in the contemporary philosophy of science of coherent scientific inference. It is thus appropriate that the discussion of the above issues should be done in the context of Bayesian probability theory.

Bayesian probability theory is an account of *de papyri* reasoning. However, as an account of the methodology of science, Bayesian probability theory is presented not only as a model of epistemological competence (that is as the principles by which inferences and deduction might be validly made in regard to a normative model of scientific reasoning), but also as a model of epistemological performance (how reasoners should go about such reasoning). This raises the question of how adequate and complete is the account provided of the actual reasoning carried out by scientists. How far can its employment by scientists (which must necessarily exist at the *de cognitio* level) be epistemological sound and reflect logical competence at the *de papyri* one? Bayesian probability has been established as an account of epistemological coherent deduction, but is this sufficient for it to also provide a model of the evaluation of theories that happens in scientific debate? Do scientists, for example, use Bayesian probability theory -- but only in the context of them also employing auxiliary epistemological quality control processes that act to ensure that this is done in a sound manner?

2.1 Bayesianism

Bayesianism is a system of reasoning in which 'subjective', and so hidden and non-public measures of degrees of belief get coherently adjusted. Bayesian theory infers the posterior probability of a hypothesis, P(h/e & k), from its prior probability P(h/k), the probability of some given data P(e/k), and the likelihood of that data given that hypothesis, P(e/h & k). (Where k is background knowledge). Thus, given the discovery of the truth of a relevant new piece of data, a Bayesian can confirm (increase the probability) or disconfirm (decrease the probability) of a theory depending upon its previously judged priors and background knowledge. What is presumed but never stated is that when humans talk about engaging in 'Bayesian inference', that the assertions that they make about probability are based upon, *and only upon,* consistent and coherent adjustment given background knowledge made through the Bayesian rule.

$$P(h/e \,\&\, k) = \frac{P(h/k)\, P(e/h \,\&\, k)}{P(e/k)}$$

2.2 Bayes' rule and valid reasoning performance

However, how far is it possible to know that when individuals assert that they are engaged in Bayesian reasoning, that they have properly created posterior probabilities from priors? If this cannot be known, then the coherence and the consistency presumption of Bayesianism can only exist at the logical *de papyri* level. If so, while it might be the case that methodologically that Bayesianism provides the foundations of valid scientific reasoning as an epistemological competence, this does not entail that it provides an account of scientific reasoning as an epistemological performance.

It seems reasonable to suggest this is the situation because it is not possible to know whether there has been a consistent creation of posterior probabilities from priors. After all, nothing in the act of individuals, such as scientists, when they claim that they have coherently obeyed Bayes' rule, actually allows them (or anyone else) to know that their degrees of belief might have not been adjusted differently (discussed below, and in Appendix 1). This situation exists because there is no factor in Bayesian reasoning that identifies whether a reasoner actually follows its rules of adjustment -- rather than merely only think and claim that is what they are doing.

For example, upon realising the truth of new evidence, a scientist when arguing for a theory might unconsciously fudge their priors as to the likelihood of this data given their background knowledge. This might happen if a scientist given evidence that at an earlier time that they would have found surprising (highly unlikely given their background knowledge), might now interpret (since they now know its truth) was not that really unexpected. Alternatively, they might slip into thinking that the new evidence has little relevance for evaluating the truth of a new theory compared to what they held when its truth was not known (when it was seen as an important consequent given background knowledge). This situation is possible due to the nonauditable nature of the background knowledge factors that construct the degrees of their belief. This nonauditability means that unless special precautions are taken, no one, not even the scientists themselves, will be aware of the existence of such post hoc adaptation of their priors and likelihoods.

This is not to overlook that the point of Bayes' rule is that reasoners should properly treat priors as priors, and reason in argument such that their degrees of belief are ruled by deductive consistency. The problem here is that the degrees of asserted belief upon which Bayes' rule operates in debate exist at the performance level, and so are not automatically determined by Bayes' rule. Due to this, while degrees of belief (or confidence) might be claimed in argument to exist in the *de papyri* realm, its actual persuasive impact upon scientific evaluation necessarily will always be done in the *de cognitio* realm. Since the latter is not necessarily transparent to the former, it follows that inferences made *de cognitio* are not necessarily valid *de papyri*.

As a result, we can conclude that while the arguments carried out by humans in the light of Bayes' rule might allow that their degrees of belief are shaped by obedience to this rule, they then again, might not, be so determined. For example, an individual scientist asserting that their arguments obey Bayes' rule, finding data that does not support their theory, can at the *de cognitio* level fudge their prior degrees of beliefs. They can do this and still maintain – albeit illegitimately their "support" for it – and there is no way (without the additional presence of practices that audit the actual consistency of claimed Bayesian reasoning) by which any scientist (including the scientist themselves) can decide whether they in fact are, or not, post hoc fudging probabilities, and so whether they are, or not, proper Bayesians.

The unexpectedness and novelty of new data, after all, is dependent upon the interpretation given it in the context of background scientific knowledge. Thus, without a means of auditing how such probabilities are constructed from such background knowledge, there is always the possibility for an after-the-fact reinterpretation whereby an individual researcher can be inconsistent and take priors to be different to what they would have done earlier. From this subtle reappreciation of probabilities, it is a short step to ad hoc modifications with theory and reasoning being adjusted so that the implications of inconvenient data can be downplayed or denied when contributed to scientific debate. Indeed, human cognition might go further, and use Bayes' rule as a guide to 'reverse engineer' 'priors' to achieve desired 'posteriors' in argument to give individual scientists – or indeed, scientists as a group, the collective illusion of having obeyed the rule. (*De cognitio* processes due to their unauditability can always masquerade themselves in this way as having been 'ruled' by *de papyri* ones.)

Scientific persuasion here, of course, to be rational, must in its logical evaluations be like justice, in that it 'should not only be done, but should manifestly and undoubtedly be seen to be done' (to quote Lord Chief Justice Hewart's famous words from *Rex v Sussex Justices*). The problem is that the coherence of Bayesian probably adjustments when there cannot be any audit of their construction from background knowledge are neither reasoner independent nor sufficiently transparent in nature to allow of their being 'manifestly and undoubtedly be seen to be done'.

At a theoretical level of epistemological analysis this does not matter (since it is concerned only with issues that exist at the *de papyri* level). But it does matter when the Bayes' reasoning is used to explain the actual methodological basis of the persuasive reasoning employed (at the *de cognitio* one) by scientists.

Here daily in the lab, in the monthly journal and at the annual conference level of exchange of reasoning, issues of "subjectivity" and probability change are particularly crucial due to the fact that degrees of belief, their adjustment, and how they are constructed from shared background knowledge are not open to public and independent verification – the 'manifestly and undoubtedly be seen to be done' factor. As a result, this lack of transparency stops any human scientist (including those asserting inferences) checking the deductive consistency of the logical steps and modifications argued or claimed to exist between priors and posteriors, and so whether there has been a fudge in Bayesian reasoning. There is always the risk that surreptitiously dressed up pieces of 'pseudo-inferences' may end up wearing the garb and language of logic –and so the imprimatur of science. Bayesian reasoning, however valid on the pages of the statistics or philosophy of science textbook (which concerns only the *de papyri* level), fails to provide the needed public auditability. This is important since it allows (if not accompanied by auxiliary processes) that humans with their limited capacities (at the *de cognitio* one) will fool themselves and others in debate that are making proper deductions, when they are not.

In consequence, if scientists followed Bayesian theorists without supplementary practices, a real possibility exists that persuasion in scientific debate would be contaminated by non-rational or even irrational processes. This would result in their evaluations creating something other than a rational investigation of the objective world. Thus, whatever its merits as an account of inference, Bayesianism cannot on its own be an account of how scientists do their epistemological argument. This is not necessarily a criticism of Bayesianism, more an identification of an epistemological incompleteness that suggests the need for further philosophical investigation. Scientists might be Bayesian reasoners but only do their reasoning in the context of supplementary epistemological practices that correct this auditability problem. This could be an interesting and fruitful area of research for philosophers of science.

2.3 Bayesianism and auditability

Not all forms of deduction, it should be noted, are effected to equal degrees by this methodological problem. For some kinds of inferences, validity can be transparent as false assertions are easily checked as to whether they are invalid. (The *de papyri* level in such circumstances is transparently accessible from the *de cognitio* one). Propositional logic is like this: if an argument is laid out in terms of explicit propositions and logical operations, any individual can translate them into symbolic notation, and verify them on paper. Alternatively this can be done by a computer and appropriate software (for example, *Mathematica*). This is because propositional deductions are constructed from entities that in their nature are overt statements and so can have their logical warrantedness checked.

However, Bayesianism is not in this way auditable. Degrees of belief (or confidence or assurance) upon which it is based are postulated states (of mind, or of knowledge). That does not mean that they cannot be translated into numerical symbol statements (of values between zero and one). But that translation process, if it is carried out, is one that is separate to Bayesian reasoning, and so not part of the Bayesian process of inference. Thus, when humans argue propositionally, they automatically do so in publicly verifiable entities, but unless humans take special precautions (discussed below) when they argue in a Bayesian manner, the validity of their reasoning cannot be known to be warranted.

A number of objections to these observations about Bayesianism with replies are presented in Appendix 1.

3 Bayesianism and the human mind

The problem of nonauditability arises because all evaluative reasoning is carried out by the human mind and so exists at the *de cognitio* level. Moreover, scientists and nonscientists use cognitive processes that are known to be innately flawed in terms of their capacity to enable logicalness. Methodological investigations of science cannot ignore this. Indeed, this is part of the interest in studying scientists and their collective reasoning. How do humans manage the trick (when they attempt to be scientists) of being rational in spite of starting out with such bad reasoning equipment?

Human thought has been established to be optimised for concerns different to those posed by logic (for example, modus tollens is both hard [Wason, 1966], and logic is argued to be of little importance to normal cognition [Sperber, Cara, and Girotto, 1995]). That logic is not immediate to our desire to reason correctly is also evidenced by history and the slowness with which humans appreciated and understood the basic principles of logic and probability. Aristotle's ideas about logic were taken as the final word for over two thousand years; however, logicians now appreciate that they are inadequate to even establish simple mathematical proofs. The work of Thomas Bayes, whose theory was published as recently as 1763 (and then only by a friend after he had died), has been largely ignored until recently. This suggests, that however much logic, consistency and Bayesian inference might be the basis of valid argument and persuasion, its principles are not so necessary or immediate to the human mind that the human mind automatically and invariably selects to use them –rather the opposite. Thus, humans when they attempt to think logically (rather than employ belief in regard to their normal concerns) face the problem of finding a way to follow the dictates of proper deduction.

Human reasoning, in particular, faces the problem that it can take itself to be confidently logical when it is not. For example, psychologists have found a broad range of reasoning pathologies exist that reflect the boundedness of human reason such as anchor effects, confirmation bias, and inadequate exploration of alternatives. Individual humans in their attempts to be reasoners, however, show chronic blindness to these limits upon the warrantedness of their reasoning. For instance, as pointed out in the *Psychological Review* by Raymond Nickerson, (1996, p. 426) ‘In the context of Bayesian decision making, people often overestimate the completeness of the sets of hypotheses provided about the possible states of the world. When given a fault tree, for example, people tend to overestimate the probability of faults listed in the tree and to be relatively insensitive to possibility not shown explicitly’. As a consequence, he notes that ‘underestimation bias’ occurs, in which ‘events not brought to mind are overlooked and the probabilities of retrieved events are distorted upward’ (p. 426).

False assertion of validity, thus, occurs even given the best intentions to be rational. Empirically, this seems to be due to people evaluating ideas in two contrasting manners (Evans, 2003; Stanovich and West, 2000). First, by using rapid, parallel and automatic implicit processes that are unconscious until their conclusions end up in consciousness. Second, by using slow sequential ones exist that allow explicit abstract hypothetical inferences (Evans, 2003; Stanovich and West, 2000). Human assertion, it seems, arises from the parallel and automatic implicit processes without adequate supervision by the slower explicit sequential and logical ones: this situation makes humans vulnerable to persuading themselves that they are logical when they are not.

Due to the capacity of the mind to proclaim itself to be “logical”, and its limited capacity to be logical, individual humans therefore readily assert, falsely – and, in spite of doing so with great confidence and often strong persuasiveness -- that they have reasoned correctly, when they have not.

This failure of logical performance to be backed by logical competence, unless rectified, would severely impair the ability of scientists to engage in persuading each other in their evaluations in an epistemologically sound manner.

Humans in their attempts to be reasoners collectively as a result need a third faculty: one that acts to guarantee the integrity that they are sound in their assertions about the logical basis of their assertions. But scientists (like all humans) are not born such an integrity faculty. However scientists are not just human reasoners but human reasoners that debate within a culture of institutions, values, and practices. Scientists therefore can arrange to create an ethos of institutional practices amongst themselves that act as epistemological quality controls upon their debate to ensure that their debate is sound.

In this use of quality control practices, they are not unique – the use of quality control practices over human interactions in regard to abstract standards is a fairly widespread in games (see Appendix 2). Though not discussed here it is also central to institutional activities such as judicial procedure, safety regulation, airworthiness certification, ancient democracy, and to a limited degree, modern western politics. Regrettably, as yet, the use of such quality control practices, though vital to the well-being of human coexistence is not amongst scholars a widely appreciated nor researched aspect of society or its institutions.

What is novel to science as a social institution – indeed defining of the nature of science as human collective enterprise and a source of its success – is the use of quality controls by its practitioners in regard to reasoning about external realities.

4 Epistemological quality control practices in science

As, it has been observed above, it is the human mind in its collective capacity to create evaluative debate, not systems of logic, that underlie rational reasoning about the world. Scientific debate, however, though it seeks to be based upon inference and evidence occurs through persuasion. This potentially makes it epistemological limited, because, as noted, there is no automatic transparency such that persuasive assertions made at the *de cognitio* level will necessarily reflect actual logical warrantedness at the *de papyri* one.

But human rationality in science seems to epistemologically succeed. This, I suggest, is due to the special nature of the evaluative debate which scientists create amongst themselves. The collective activity of science, by design and intent, aims to generates an evaluate debate under the governance of epistemological practices that ruthlessly ensures the soundness of its evaluations. It is these practices that link the logic of the *de papyri* level with the assertions made in scientific debate at the *de cognitio* one.

This raises the question as to the nature of the quality control practices that scientists use to ensure such soundness. As noted proportional logic can be automatically checked so that assertions made at the *de cognitio* level can be known directly to be valid at the *de papyri* one. Unfortunately, as noted, much scientific reasoning is probabilistic (if, as seems likely, it is well described by Bayesian inference) and so depends upon hidden coherence, and as a result is not overtly verifiable.

Quality control practices can achieve epistemological soundness in scientific debate both in negative and positive ways (though in reality they are closely intermixed).

- Negatively –by blocking the influences of what might be called the "pathologies of evaluation". These cause factors other than truth and deduction to compete against them in theory evaluation. Negative quality controls by inhibiting their presence in debate increases the independence of a theory evaluation from such human cognitive weaknesses, and so the ease with which it is compromised by them.
- Positively –by increasing the transparency of assertions to the *de papyri* valid inferences upon which scientists claim that their persuasiveness is founded.

4.1 Negative Acting Quality Controls

As noted above, people's reasoning is affected by diverse biases. These include not only the cognitive biases already mentioned but also social ones. For example:

(a) In general people seek to hold the ideas of their social group. Indeed for people to form social groups to which they feel a sense of belonging, human social psychologically seems to require them holding ideas in common, and engaging in activities which display these shared beliefs.

(b) People seek to be deferential in their beliefs by holding only those ideas approved by those with a higher status and therefore social rank over them. As with the formation of social groups, there appears to be a social psychological process by which the holding of ideas upholds dominance relationships. The social position of individuals, for instance, is threaten if subordinates take it upon themselves to accept as true ideas that conflict with those of their superiors.

These social disruptions of theory evaluation need to be quality controlled out of the scientific evaluative debate if false ideas are not to be mistakenly propagated. This is because they add factors that compete against the selection and preference of theories in regard to their rational soundness.

In general this is done by the egalitarianism and social tolerance ethic in science. This ethic is helped by the fact that most scientists spend part of their academic career in countries and often cultures other than the one in which they were raised, or work with colleagues that are such visitors. Further, the communication of science is usually done as far as possible in a way that is blind to the status and cultural background of individuals. For example, most major science journals print on their papers only the name of an author without indicating whether they hold an important position such as a head of institute, or not, such as a new researcher, or even a student. Blinding submitted papers for peer review is also common: this not only removes the factor of bias against known individuals but also biases linked to judgements about the author's institutional and social status. There is minimal honorific language in science though this is culturally widespread in many societies.

Another factor is the taboo against opinions and beliefs due to the personal and social emotional investment that associates with them. Scientists to avoid nonrational factors in general seek to hold as few personal opinions about scientific ideas as possible, and where they hold them, to be aware that they are opinions and so likely to be constructed in regard to nonrational biases. The general ethos of science is that assertions should not be derived from opinions, but opinions derived from demonstratable facts and inferences and so depend upon those facts and inferences, rather than any incidental personal or other emotional investment in them as beliefs.

4.2 Positive Acting Quality Controls

Minimizing social, affective and cognitive pathologies upon theory preference is insufficient to link the *de cognitio* and *de papyri* levels – there also needs to be positive actions. One way of doing this is by exploiting the logical asymmetry between confirmation and refutation (this is discussed in Appendix 3). Another way is to increase the transparency of claims at the *de cognitio* level to the *de papyri* one by externalising the background knowledge factors that shape and construct prior probabilities, likelihoods and other degrees of belief and confidence. This makes their adjustment in a post hoc manner difficult without this becoming apparent, since their basis is publicly available to others and so auditable to post hoc change.

There are several ways in which such transparency can be maximized.

4.2.1 Publicness

Every factor which shapes probabilities, likelihoods and degree of beliefs and confidences should be -- in so far as it can be -- arranged be made accessible and checkable by other scientists. Science thus needs to be driven by the ethos that what matters is always what is open and above board. Scientists as a result of these values, strongly reject, for example, hidden experiments, and personal sources of evidence, as a basis of theory support. In general, the only evidence that carries evaluative weight is that which can be seen to be autonomous from personal interests, biases and other nonlogical factors. Scientists, moreover, to this end have a responsibility to make as explicit as they can the procedures of their experimentation, statistical and data evaluation. Scientists have to explain the nature of their theories, why they are important, and how they fit in or not with other already accepted ideas. Scientists which fail to do this gain less standing -- reputation -- amongst their peers than those that actively make what can be made public, open and explicit.

4.2.2 Clarity

Auditing is impossible without clarity. If the factors that shape degrees of belief are to be independent of particular individual scientists, then they must be communicable. By definition, any belief held by an individual that is opaque is not communicable and so cannot to be judged independently of them. Therefore, scientific communication should aim to be clear. This again is a factor that determines a scientist's reputation amongst their peers.

4.2.3 Open debate

The background knowledge used to shape probabilities, likelihoods and degrees of belief needs to be articulated. The first attempts of any individual to specify what is relevant background information to any probability will not necessary capture its most pertinent connections. Therefore the connections between ideas and hypotheses will often need to be challenged to clarify and remove ambiguities, inconsistencies and incompletenesses as to how they might be interpreted. Scientists for that reason must engage in discourse and debate that spells out exactly what they are asserting, and what matters to their construction of probabilities so that others can follow how background knowledge shapes their inferences.

Particularly at the level of conferences and forums, there is a reputation effect. Certain conferences (and their organisers) gain standing because they are organized in an effective way to elicit such debate. There are many skills here that are not fully appreciated and fail to receive a proper investigation. At a conference, for example, there might be five minutes allocated for discussion at the end of each talk, and with session of talks, half an hour of general discussion, and at the end of the conference itself, an afternoon of free debate. The person in the chair can make considerable difference to quality of the debate that then happens. It is a vital part of the reason for attending confidences, yet there is a surprising lack of investigation into the skills needed if this is to be effective. One might make similar comments about editors and the organization of reviewers, commentaries and letter-to-the-editor in journals.

4.2.4 Criticism

While individual humans may be sometimes self-critical, in general, they are often, and indeed, usually not. Thus, it is more likely that mistakes in theories, poor specification of their contents and their links to background knowledge will be pointed out by others not personally involved, than those that are so committed, and so without career, or self-esteem investment in them. Science, thus, seeks to promote an ethos in which scientists expose their ideas to the detection of error, ambiguity and poor articulation.

Criticism is closely linked to open debate. It is not enough that opportunities to debate exist: people must seek to create ideas that they then expose to challenge, and individuals must exist that are prepared to criticise them. This involves not only a willingness to offer ideas and criticism but also an ethos that makes it safe. At a conference, a PhD student giving their first talk receives (or should receive) a very different reception to that given a named chair professor. Criticism has to be tailored so that people are encouraged to offer new ideas and work to improve their full exposition. A new participate in science will often make many errors due to unfamiliarity and anxiety. A responsible critic will make comments that encourage them.

Indeed, this applies to all criticism. Many journals, for example, ask that reviewers view their task not as determining the acceptability of a submission but as offering ways of improving an submission irrespective of whether it should be accepted or not.

For example consider this statement about the reviewing policy of the journal *Cognition* (Altmann, 2007: p. 3.).

> Reviewers are reminded that they have a responsibility both to the science and to the authors who are trying to advance that science. This responsibility includes helping authors better their papers and, if necessary, better their science. The role of the editors and the reviewers is as much to reach a consensus on how the author could improve the impact of their research as it is to reach a consensus on whether a paper should be accepted, sent back for revision, or rejected. And if reviewers can suggest alternative, more accessible, ways of 'packaging' those hypotheses, or of packaging the data themselves, the review process can better serve the authors. In this regard, reviewers are explicitly asked to consider their role as being more akin to a mentor than to an examiner.

There is a whole ethos hidden away in the carrying of science that seeks to minimize ad hominem remarks and ensure criticism are always constructive. The reputation of many a scientist is built not just on their work but their ability to identify and comment upon the weaknesses of other's work in such a manner that it is not felt as a negative criticism by those in the audience and those at the receiving end.

4.2.5 Conjectures

Another important part of degree of belief is the presence or absence of theorized alternatives. As noted above, probabilities depend upon the adequacy of the exploration of different possibilities. A high degree of belief in a theory, moreover, will often be held due to the lack of, or the suppression of, ideas that compete with that theory. Moreover, it is only in the context of alternatives that a theory can be fully detailed and theoretically expounded. Therefore, theory status needs to be constantly challenged by possible alternatives. This requires the seeking and the valuing of the promotion of conjectures which challenge received ideas. Scientists should aim to maximise the space in which theories get examined so they can see how data might or might not be explained by them. This again links closely with open debate and criticism. There is no point in open debate, if no one offers alternatives to time-honoured ideas.

While, alternatives must be encouraged, they cannot of any kind. There is at the edge of science, many advocates of strange and wonderful ideas that if taken seriously would take time and resources from the debate of better ideas. It is often very hard, however, to know what are potentially profitable conjectures from those that are cranky and screwball. The history of science is full of examples where establishments stopped the debate for long periods of ideas (i.e. plate tectonics) that later turned out to be important not idiosyncratic. However, against this, the dismissal of many ideas has been shown has been in retrospect to be fully justified as their inappropriateness has only increased with the clarity of time. But equally there are cases where new ideas were readily and correctly accepted (Linus Pauling's instant acceptance of Crick and Watson's double helix model of DNA, for example), and where establishment outsiders were treated with appropriate encouragement (Max Planck's associate editorship of *Annalen der Physik* and the *annus mirabilis* 1905 papers of patent clerk "third class" Albert Einstein, are an out standing example here).

4.2.6 Reputation

As noted science works to a considerable degree through the reputations that scientists gain in how they participate in scientific debate. Part of the ethos of science is that scientists hold – or should hold -- strong prejudges against those that seek to propagate their beliefs in ways that bypass the above quality controls and so sound evaluation.

Cheating, for example, destroys a scientific reputation. No just in the narrow sense of faking data but discourse disingenuousness in any form. Spin, PR and bluff might be acceptable in contemporary politicians but amongst scientists engaging in such dubious argument tactics stops their peers taking them and their work seriously.

Indeed, a scientist that engages in disingenuousness in regard to quality controls risks ceasing to be considered as a scientist. Science lives under the shadow of pseudoscience -- that is ideas advocated to appear to have the imprimatur of science without its substance. The main distinguishing feature of pseudoscience is the disingenuousness with which it evades evaluation done under scientific quality control practices. The separation of pseudoscience from science is not one of logic but one of the acceptance of the quality control practices, and so soundness in its evaluation by its advocates.

In contrast, scientists that honestly acknowledge mistakes and failings that are nonintentional do not get penalised. Part of science is not punishing those that go along research avenues that turn out to be without fruit. At conferences, audiences have been known to applaud speakers that accept a theory long held by them is not worth -- in view of new data -- further investigation. Science is not only about advocating bold conjectures but dropping ones they have ease to merit more backing.

None of the above quality control practices touches upon whether a theory is true or not. A theory might be true irresponsive of how it is evaluated. There is no logical reason, for example, why a theory propagated due to a dictator's diktat without open debate, clarity, criticism or competing alternative conjectures need necessarily be false. Our intuitions might be that in such circumstances it is unlikely to be true. But the issue here is not such intuitions but the nature of soundness. Whether a theory is true or false does not determine its soundness since soundness is an issue of how it is evaluated. A theory which is unsoundly evaluated may be true, the problem is no one can know whether everyone's confidence that it is true is warranted or not. Soundness is not a property of ideas but the nature of the evaluation to which they are exposed. It is a property created by the situation that ideas are evaluated by fallible humans – the problem of humans evaluating in the *de cognitio* world, not in the *de papyri* one in which validity exists. To bridge between them, scientists engage in the quality control practices such as those mentioned above. It is these that bring the validity of assertions made at the *de cognitio* level as close as is possible by human minds to their claimed basis in logic. It is due to quality control practices that scientists can reason effectively with logical deduction that exists at the *de papyri* level, even though they, and their reasoning, necessarily happens in the *de cognitio* one.

An analogy exists here with artefacts. Humans are weak in strength and perception but can act beyond their natural abilities with the assistance of machines (hoists and bulldozers) and optical devices (microscopes and telescopes). Science in quality control practices provides the human mind with a means to overcome the situation that it is frail in its innate reasoning capacities.

5 Quality control practices and Popper

It hardly needs to be observed that the quality control practices noted above contain many of the ideas and principles that Popper identified with scientific rationality. Popper, however, attributed their importance to the making of scientific valid inferences (the *de papyri* level). Here this paper attributes them instead to a different epistemological realm – the *de cognitio* level and the need of scientists to ensure scientific soundness.

This shift might explain Popper's paradoxical better reception by scientists than by philosophers of science. As Popper famously noted, 'Here I am being showered with honours as no professional philosopher before me; yet three generations of professional philosophers know nothing about my work' (Bartley, 1982, p. 272). Working scientists, unlike philosophers of science, face daily the problem posed by their reasoning being epistemologically compromised when they engage in evaluative debate. They are thus acutely aware that accounts of scientific reasoning should explain how its inferences can be made sound at the *de cognitio* level. Those professionally studying scientific rationality, in contrast, have no such immediate need to be aware of the problem of epistemological soundness as they tend to work at the *de papyri* level of logical competence.

This might also account for the absence of philosophical interest in epistemological quality control practices: lacking an immediate need that forces them to be aware of the importance of soundness, academics studying reasoning focus nearly exclusively upon the paper warrantedness of scientific ideas. This has misled them to overlook the essential role in scientific persuasion of quality controls. Scientists, in contrast, have found it necessary from daily experience in evaluating theories to use quality control practices to effectively minimise the impact and disruption of cognitive biases and false claims.

Philosophers might argue against extending scientific methodology to include quality controls and soundness by arguing that they deal with an aspect of science – the principles of its logical (or *de papyri*) rationality – that can be isolated from its actual employment by scientists (*de cognitio*) in debate, persuasion and argument. Philosophers, in this view, have no interest in how scientists actually reason (they are not social anthropologists as to how scientists actually work at the lab bench) but instead are interested, and only interested, in the logical competence of argument needed to establish valid scientific conclusions.

However, the issues of soundness, and the validity of scientific rationality are not methodologically separable: the logic that backs the rationality of science requires that what is claimed to be actual, is, in fact, actually what is asserted. Moreover, the philosophy of scientific reasoning is about reasoning done by scientists, not abstract processes disembodied from actual reasoners that exist only upon paper. (Many methodologists, in the course of their work, indeed frequently cite evidence about how scientists reason in regard to such things as ad hocness, simplicity and so on.) Understanding this rationality is therefore not an "anthropology": it is concerned with how scientists should go about exercising their reasoning faculties so that they actually reflect logical principles. Consider, the analogy of car driving: one might study how car drivers actually drive (an anthropology of car driving), but one might also study the principles that should determine the 'Highway Code' that lays down how they should drive if they are to do so safely (the normative requirements for safe car driving). What is proposed here is the later kind of study except, it concerns the requirements for how scientists should reason if they are to make, rather than just think they make, rational inferences and theory selections (for further arguments, see Appendix 4).

6 References (including those cited in the following appendixes)

Altmann, G. (2007). Editorial: Journal policies and procedures. *Cognition, 102,* 1–6

Bartley, W.W. III (1982). A Popperian harvest. In P. Levinson (Ed.), *In pursuit of truth*. pp 249-289, (Humanities Press. New York, 1982).

Evans, J.S.B.T. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Science, 7,* 454-459.

FIDE. (2005). *FIDE handbook. Section E.I. laws of chess*. Available at http://www.fide.com/official/handbook/E_I_01a.pdf

Geisler W.S. and Kersten D. (2002). Illusions, perception and Bayes. *Nature Neuroscience, 5,* 508-10.

Howson, C. and Urbach, P. (1993). *Scientific reasoning: The Bayesian approach (Second Edition)*. Chicago: Open Court.

Jaynes, E.T. (1988). *Probability theory: The logic of science*. (available on the internet at http://bayes.wustl.edu/etj/prob.html)

Nickerson, R.S. (1996). Ambiguities and unstated assumptions in probabilistic reasoning. *Psychological Review, 120,* 410-433.

Sperber, D, Cara, N. and Girotto, V. (1995). Relevance theory explains the selection task, *Cognition, 57,* 31-95.

Stanovich, K.E. and West, R.F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavior and Brain Science, 23*, 645–726

Wason, P. C. (1966). Reasoning. In B. Foss (Eds.), *New horizons in psychology* (pp. 135-151). Harmonsworth, England: Penguin:

Weiss Y, Simoncelli E.P. and Adelson EH. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5,* 598-604.

World Bridge Federation. (1997). *Laws of duplicate contract bridge*. Available at http://www.worldbridge.org/departments/laws/laws.htm

APPENDIX 1

(a) The invalid-breaking-of-rules argument

Bayesian theorists may object that the possibility that humans might engage in improper or masquerading inference in no way undermines Bayesianism as a model of scientific reasoning. After all, a scientist that upon finding data that goes against a theory, then retrospectively adjusts its prior probabilities to protect it, or that allows that their subjective probabilities to be adjusted inconsistently by psychological factors like confirmation bias, and selective attention to alternatives, is doing something other than Bayesian reasoning.

Such theorists might cite the example of chess. Here, the possibility that humans might make invalid moves, after all, does not stop the existence of chess tournaments. But this objection ignores that chess is a system of rules and practices that is intrinsically sound because it is an observed board game and ruled by procedures that stop cheating. As a result, what is claimed to take place is easily inspected as having actually taken place – moves are done publicly, and thus verifiable. Not only do rules exist that govern the theoretically permissive moves of chess pieces but also practices exist that ensure that human chess players abide by these rules: for example, a player, once they have touched it with his or her hand, is required to move a piece (see appendix 2 (b)). That the pieces and their moves are public allows an opponent to analyse whether they have been played according to the rules and so puts them and others in a position to detect intentional and unintentional errors. The play of chess is therefore more than the set of theoretically permitted moves since it also includes practices that makes manifest the physical honest playing of them. As a consequence, its play happens in a manner that is transparent and auditable to the theoretical one of its rules, and so has in regard to them verifiable integrity.

For this reason, the argument based on the invalid-breaking-of-rules cannot be accepted: nothing in (or outside) Bayesian reasoning, except in special circumstances (such as those enforced by the above discussed quality control practices), methodologically acts to enable scientists to distinguish between real Bayesian reasoning and conscious or unconscious created 'Bayesian' pseudo-reasoning. Without such a means of detection, and the lack of necessary transparency between the *de papyri* and the *de cognitio* levels, methodologically it can never be known whether scientists on any particular occasion actually reasons validly in a Bayesianism way, rather than existing only in the state of believing themselves to reason in such a manner (while in actuality, persuading and holding ideas upon whimsy and capricious prejudice). The chess analogy with Bayesian theory would be chess done between players (perhaps having taken a drug that impairs short-term visual memory) that are blind and amnesic to the legitimacy or otherwise of earlier moves made by their opponents. Such chess playing would radically change the nature of chess – even though it would not stop in principle the competence and capacity to play chess – since without the transparency and so the cognitive ability to check for correct play, its realisation as a performance would be radically changed – no player could know whether an opponent won through good play, or by cheat. There might be chess wins and "chess wins" but no way of ascertaining the difference between them.

(b) Ceasing to be Bayesian argument

Another related objection is that scientists should not engage consciously or unconsciously in adjusting probabilities to fit the conclusions they covet: the rules of

Bayesianism reasoning exist to be obeyed, or scientists simply cease to reason in a Bayesian manner. But merely stating that scientists should act in a particular way is not a satisfactory answer since it does not tell scientists how they can know when they have engaged properly in Bayesian degree of belief adjustment. The logical competence to engage in Bayesian reasoning given by Bayes' rule does not lead by any auditable means at the performance level to information about the soundness of that performance. It may be, for example, that human cognitions are so limited that humans in their attempt to reason cannot suppress the unconscious processes that surreptitiously adjust subjective probabilities, however, much they understand, value and seek to practice correct Bayesian reasoning. How can it be philosophically established that our reasoning does not exist in such a deductive topsy-turvy world?

(c) Bayesian theory is 'impersonal and objective'

Howson and Urbach try and minimise the methodological disadvantage of subjectivity by claiming that Bayesian theory is 'impersonal and objective' (1993, p. 419). But it is so only in a very special sense of the meaning of these words. As the context of their statement makes clear: 'As far as the canons of correct inference are concerned, neither logic [deductive or Bayesian] allows freedom to individual discretion: both are quite *impersonal* and objective'. But only at the *de papyri* level can we assume that the canons of correct inference rule, and that there is no "individual discretion". Leave that theoretical world where it can be specified that things happen as required on paper, and the question whether inference has been correctly done, or "individual discretion" in a tacit way has been exercised becomes important. Philosophers of science, after all, are not interested in the *de papyri* level by itself: they are concerned with the methodology of science in which actual scientists exploit the principles of logic to enable them to rationally evaluate theories. At this level, while the deductions of propositional logic are impersonal and objective (their *de cognitio* realisation makes their *de papyri* status transparent), those of Bayesian deduction are not. Howson and Urbach, in fact, acknowledge this in their words, 'impersonal' and 'objective': in the *de papyri* world, there is only reason and no individual reasoners, and so no issues of whether deduction is done without "individual discretion", or not. Moreover, the book from which I get this quotation, is called, after all, *scientific reasoning*, not *paper logic reasoning*). The analogy Howson and Urbach seek to propose between the propositional and Bayesian deduction is therefore inappropriate.

(d) Fails in long term argument

Bayesians might further argue that its inferences are not hidden, or at least, do not need to be hidden. Individuals that claim to – but do not – engage in Bayesian reasoning will contribute to an activity that will show no long-term epistemological progress. They will not make valid deductions, and as a result, they will not improve their theories. In the long run, this lack of advancement will be noticeable. But, of course, this absence of success (even if it could be detected -- which would involve a bit of a paradox), provides no evidence about the consistency of adjustment in any particular case of reasoning. Lying may be fruitless, but we would prefer a direct means to know whether people are or are not telling fibs than the long-run effects of false assertions.

(e) Probabilities can be made objective

Bayesians might also argue that the auditability problem is easily corrected. Degrees of belief can in theory be made reasoner independent. This happens when they are transformed into numerical values. Reasoners could, if need be, when claiming to make Bayesian inferences therefore layout the steps of their reasoning, and so make the consistency of their deductions auditable to others. We can imagine a science where scientists banked their probability judgements in depositories specially set up to record priors, and then did their research; then depending upon its results openly engaged in a public adjustment of these probabilities. Thus, there is nothing in principle why Bayesian reasoning should be liable to surreptitious nonreasoning processes with people asserting themselves to be Bayesian when they are not. The problem here is that no important scientific discovery has been based upon such banked probabilities, nor is there in present science any means for scientists to do this. There is no public depository of probability judgements; no important journal publishes them, and few scientists, even if they were recorded, would take note of them. That this is not the practice of science is curious given Bayesian reasoning is supposed to underlie its rationality. Either scientists do not reason in a Bayesian manner, or they employ it in a way different to that suggested at present by Bayesian orientated philosophers. Further, it may be observed that such a science would involve separate, additional and potentially problematic processes that are not part of the Bayesian theory.

(f) The visual system and Bayesian theory

It is now a commonplace in neuroscience to call the brain a "bayesian machine". The visual system is well modelled, for example, in terms of Bayesian processes (Geisler & Kersten, 2002; Weiss, Simoncelli & Adelson, 2002). Such models, for instance, can explain several optical illusions and many aspects of human reasoning (see for example the articles in the special issue of Trends in Cognitive Sciences Vol.10 No.7 July 2006). It may be argued that human reasons likewise should be equally well modelable in terms of Bayesian processes. The reply here is that this paper does not deny that human reasoning should not be modelled in terms of Bayesian processes. This paper argues instead that successful Bayesian reasoning for its existence depends upon additional auxiliary extralogical processes to create soundness. There is no reason to assume that this is not also the case in the brain with the contribution of Bayesian processes to sight and other cognitions. The neural interactions, for example, that generate Bayesian processes might designed by evolution to be sound from corruption in how they handle information coherence and the neural equivalents of degrees of belief. In understand this, I suggest, there is a parallel between the foundations of neurological processes and those studied by philosophers of science: both need to account for how Bayesian information processing can be done on processes that are not Bayesian and perhaps even antagonistic to it.

APPENDIX 2

(a) Rules and quality control practices

Scientists are not alone in using quality controls. Indeed, avoiding falsity or defectiveness masquerading as "validity" or "functional" is a common problem throughout human affairs.

It is, for example, the problem of "is what is on the label in the box?" It is a property of complex manufactured goods such as electrical ones that defective ones resemble in all respects functional ones – at least prior to their being taken out of their box and used. Chemical measurements also have this property: a false reading about a pollutant cannot be distinguished without further independent tests from a true one. In both areas, concern exists with "quality control" – that is with carrying out procedures that ensure that what might appear to be a valid working electrical goods or pollution readings are actually what they claim to be. Such quality control concerns the development and use of practices that ensure that what is on the label can be verified to be in the box: that the DVD player works to its specifications, rather than merely looking like one, or that the reported reading actually reflects the contaminates in a sample rather than only appearing to do so.

Scientists are thus not alone is using quality verification practices. Electrical goods and pollution readings, on the other hand, lack a direct parallel with the two *de papyri* and *de cognitio* levels of logical reasoning.

However, a relationship exists with them (though in an nonepistemological context) in games and their rules and practices. Games such as chess or contract bridge like scientific inference exist both at *de papyri* (rules) and *de cognitio* (play) levels, and can contain potential "falsity" which such rules and practices "quality control" seek to inhibit or prevent.

(b) Chess

International tournament chess and contract bridge lay down regulations for matches. Some of them are rules in that they describe the game and others are practices in that describe how they game is played. The *International Chess Federation Handbook* (FIDE, 2004), for example lays down that:

> 3.2 The bishop may move to any square along a diagonal on which it stands.

But it also states that.

> 4.1 Each move must be made with one hand only.

Intuitively the first regulation is of a somewhat different nature to the second, and parallel in this to the *de papyri/ de cognitio* difference argued to exist above in logic. A game of chess without regulation 3.2 would not be chess as we understand it, but we could readily image a game of chess with a regulation different to regulation 4.1. that was still chess. For instance, it is difficult to see how rule 4.1. can apply to computer chess with keyboard input, and exceptions must apply to those lacking hands who use their toes or mouths. Regulation 3.2. is a rule, while 4.1. is a practice.

Regulation 4.1 exists because humans cheat intentionally or otherwise. Not only does it help to stop cheating but as importantly prevents the suspicion that cheating might have happened. It describes a quality control practice that creates transparency so that an actual physical game of chess can be seen to follow the abstract *de papyri* rules laid down for the playing of chess. Regulation 3.2 in contrast has nothing to do with moves that might have been made by a slight of hand: it defines what is a right and wrong move. It is a rule.

(c) Contract Bridge

As in chess, the international organization covering contract bridge tournaments lays down regulations (World Bridge Federation, 1997) (in this case called "laws") for contract bridge which divide into rules and practices. For example the first Law defines what a pack of playing cards consists of.

> Law 1. Duplicate Contract Bridge is played with a pack of 52 cards, consisting of 13 cards in each of four suits. The suits rank downward in the order spades (♠), hearts (♥), diamonds (♦), clubs (♣). The Cards of each suit rank downward in the order Ace, King, Queen, Jack, 10, 9, 8, 7, 6, 5, 4, 3, 2.

This "law" is a rule. It defines in part what makes a game of cards, a game of contract bridge. If this rule was to be different, for example, if it allowed two packs, or did not allow one of the suits, then the game of contract bridge would be radically different in nature. It is the *de papyri* level of contract bridge. In the collection of "laws", however, are regulations that are better considered as practices that exist at the *de cognitio* level to enable the game to be engaged in by human players, for example, "law" 73 contains this subregulation.

> Law 73. A. 2. Calls and plays should be made without special emphasis, mannerism or inflection, and without undue hesitation or haste.

We can imagine contract bridge being played without this rule or a very similar version of it. For example, in some tournaments calls are not made verbally but given through the use of bid boxes and players are hidden from each other by screens. This regulation does not define the game but ensures that the rules that define it are effected in actual play – it is a quality control. A player that communicated to their partner information about their cards by how they made a call would put them at an advantage to their opponent players.

In both chess and bridge, the functional difference between rules and quality control practices are mixed up in the regulations that govern the playing of these games. This probably reflects no more than a lack of concern for the distinction amongst the players of these games. They are interested, after all, to play their games effectively, not philosophise about the methodology of the regulation used to organize how they are played. But it could also reflect a conceptual limitation: there seems to be a general failure to recognize in human society this difference in these and other activities. There is, after all, as far as I can establish no awareness of this distinction in the analysis of science.

APPENDIX 3

Selective Bayesianism

Scientists can increase the auditability of Bayesian reasoning in a further way that needs to be noted. Auditability can be enhanced by limiting deduction to those cases and circumstances where priors can be adjusted free of the possibility of post hoc fudging. An implicit assumption in Bayesian methodology is that where a probability inference can be made, it should be made. But there is no obligation in the application of Bayesian theory to reasoning to require that all opportunities to adjust prior probabilities should be exploited to create new posterior ones. People could decide to reason with a limited use of Bayes' rules in which adjustments only take place when they are sufficiently public for them to be audited. This might seem epistemologically inefficient: but it would be false efficiency –if deduction was done in a manner that allowed people to making inferences such that they were taken to be valid when they were not. In other words, throwing away information can be methodologically rational if the alternative is nonrationality.

Many Bayesians might not call such reasoning Bayesian, since they implicitly and even explicitly take it that all relevant information must be used to adjust priors. For example, Edward Jaynes, 'The essence of ``honesty'' or ``objectivity'' demands that we take into account all the evidence we have, not just some arbitrarily chosen subset of it. Any such choice would amount either to ignoring evidence that we have, or presuming evidence that we do not have' (1988, p. 402). But there is no obligation to reason efficiently with all the information that is available, if we feel that we are unable to properly process it all. After all, if we lack the opportunity to audit certain kinds of deduction, we can decide not to include them in public argument. This might slow our ability to find the truth from what might be a faster route in an ideal world. But not living in an ideal world, we must reason in such way that the fudging of degrees of belief cannot take place – auditability matters to rationality much more than informational efficiency. One objection Bayesians might raise is that any selectivity in the use of the rule might bias theory selection. However, what is proposed here is a methodological restriction that would exist prior to theory evaluation, and moreover, would exist at a public level. The application of such a restriction is thus objective and so auditable. Scientists, I suggest, therefore might use Bayes' rule, but only in a cut down manner.

Bayes' rule offers one strong public inference that can underlie such auditability: post hoc revision is detectable when a theory is false given some fact i.e. $P(H|D) = 0$, and D is found to be true (as noted by Howson and Urbach, 1993, p. 119). Refutation is nearly always an auditable inference, confirmation rarely, if ever, is. Therefore, scientists should seek where ever possible to reason using this limited case of Bayesian inference. If a theory entails a false consequence then it must be false: we must always seek to refute. Scientists to avoid surreptitious fudging can obey this aspect of Bayes' rule, and ignore its potential for confirmations on the grounds that such belief changes are unauditable.

APPENDIX 4

The philosophical need to study this problem

Much discussion of scientific methodology ignores the *de papyri/ de cognitio* distinction and treats validity in the former as automatically entailing warrantedness in the latter. As a result, such investigation is not likely to offer an effective, complete or fully insightful analyse of sound reasoning, particularly as done by scientists since there is no automatic entailment of such warrantedness. *De papyri* reasoning concerns only logical competence, that is the ability of facts and deductions to make valid conclusions. To do this, *de papyri* logic is, and only is, a paper logic: that is, it is a process solely concerned with the abstract manipulation of symbols. *De papyri* principles of inference thus only deals with the capacity of such abstract entities to lead to valid deductions. As such, it says nothing about the epistemological principles by which humans in their attempts to reason can ensure they do so in a logical manner: that how in their performance as reasoners, they might actually realise these logical capacities. This is a problem since nothing in the nature of logic stops people "passing off" nonlogical assertions as logical ones.

Another way of putting this issue is that methodologically scientific assertions, in spite of appearances and claims made on their behalf do not automatically, or even necessarily, reflect the abstract logical entities they profess to manipulate. Assertions (whatever their pretence to logic) are merely words that have no direct link with logic. Humans use words, after all, to gain the status and persuasiveness: it is easy to suggest logical validity to gain approval, when in fact what is asserted is not backed by valid deduction. Human nature thus ensures that most assertions of logical validity are only that – assertions.

Since scientific methodology addresses the analysis of scientific reasoning both at the levels of logic (epistemological competence) and of how logic is realized (epistemological performance), philosophers of science need to investigate the methodology of both, and so the processes by which humans can rationally use facts and deductions to come to valid conclusions in spite of the limits of humans as reasoners.

This is not to deny that all aspects of reasoning done by actual scientists will always be of methodological concern to philosophers of science: the economic aspects of the processes by which science journals are edited or conferences are organized, for example, would seem to be irrelevant to philosophers. But philosophers cannot isolate their discussions of rationality and scientific methodology from the practices by which actual scientists seek to reason well in journals and conferences. The interest in understanding scientific knowledge, after all, derives much of its appeal from the success of past scientific reasoning. The study of science as a result must include a concern with the nature of the actual (and presumably not unsuccessful) attempts to reason validly about the world. Thus, philosophers must have a concern with the level of its *de cognitio* practices. Philosophers of science cannot discuss actual cases of scientific knowledge, and then turn around and assert that their concerns are confined only to theoretical reasoners. By having discussed actual cases of successful science, they have shown a concern not only with hypothetical processes, but with the products of actual and effective scientific reasoning, and so with that the methodological situation of real life scientific reasoning. The methodology of science, therefore, must be concerned about its rational performance as much as its rational competence.